

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 00:09:27 ; Search time 1962.22 Seconds
(without alignments)
17301.659 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatccgcgaaggtcacaca.....ccataagagtttgatgcac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 5735 | 100.0 | 5735 | 12 | Adi28886 Inducible |
| 2 | 5045 | 88.0 | 5443 | 5 | Aad02408 Mouse alp |
| 3 | 5045 | 88.0 | 5443 | 6 | Abi49724 Mouse pro |
| 4 | 1279.8 | 22.3 | 1679 | 2 | Aat31006 Mouse car |
| 5 | 1279.8 | 22.3 | 1679 | 6 | Abv74357 Mouse car |
| 6 | 1279.8 | 22.3 | 1679 | 12 | Adl14241 Mouse car |
| 7 | 407.6 | 7.1 | 5190 | 13 | Adg97705 Rabbit al |
| 8 | 311.8 | 5.4 | 520 | 2 | Aat11359 Bidirecti |
| 9 | 311.8 | 5.4 | 520 | 2 | Aat45722 Bi-direct |
| 10 | 311.8 | 5.4 | 520 | 2 | Aav60079 Bidirecti |
| 11 | 311.8 | 5.4 | 520 | 2 | Aax60045 Bidirecti |
| 12 | 311.8 | 5.4 | 520 | 3 | Aaz56126 Bidirecti |
| 13 | 311.8 | 5.4 | 520 | 4 | Aah47631 Nucleotid |
| 14 | 311.8 | 5.4 | 520 | 5 | Aah25572 Nucleotid |
| 15 | 311.8 | 5.4 | 520 | 8 | Ac94734 Tet coord |
| 16 | 311.8 | 5.4 | 520 | 8 | Abx16487 Bi-direct |
| 17 | 311 | 5.4 | 4438 | 10 | Aad63234 ptetO7Sag |
| 18 | 311 | 5.4 | 4479 | 10 | Aad63233 ptetO7Sag |
| 19 | 311 | 5.4 | 4556 | 10 | Aad63232 ptetO7Sag |
| 20 | 311 | 5.4 | 4556 | 11 | Aad59258 ptetO7-Sa |

| | | | | | | |
|------|-------|-----|------|----|----------|-----------|
| C 21 | 311 | 5.4 | 6346 | 10 | AAD63231 | ptetO7Sag |
| C 22 | 311 | 5.4 | 6346 | 11 | ADOS9257 | ptetO7-Sa |
| C 23 | 311 | 5.4 | 6423 | 10 | AAD63230 | ptetO7Sag |
| C 24 | 311 | 5.4 | 6424 | 11 | ADOS9253 | ptetO7-Sa |
| C 25 | 311 | 5.4 | 8287 | 10 | AAD63235 | ptetO7Sag |
| C 26 | 311 | 5.4 | 8364 | 10 | AAD63236 | ptetO7Sag |
| C 27 | 310.4 | 5.4 | 450 | 2 | AAT06870 | PhCMV*-2 |
| C 28 | 310.4 | 5.4 | 450 | 2 | AAQ76267 | PhCMV*-2 |
| C 29 | 310.4 | 5.4 | 450 | 2 | AAT11356 | Minimal C |
| C 30 | 310.4 | 5.4 | 450 | 2 | AAV60081 | Cytomegal |
| C 31 | 310.4 | 5.4 | 450 | 2 | AAZ27903 | PhCMV*-2 |
| C 32 | 310.4 | 5.4 | 450 | 2 | AAX81722 | A human c |
| C 33 | 310.4 | 5.4 | 450 | 2 | AAX60047 | Minimal H |
| C 34 | 310.4 | 5.4 | 450 | 2 | AAO11367 | PhCMV*-2 |
| C 35 | 310.4 | 5.4 | 450 | 3 | AAZ56128 | Tetracycl |
| C 36 | 310.4 | 5.4 | 450 | 4 | AAO09835 | Human cyt |
| C 37 | 310.4 | 5.4 | 450 | 4 | AAH47633 | Nucleotid |
| C 38 | 310.4 | 5.4 | 450 | 5 | AAH25574 | Nucleotid |
| C 39 | 310.4 | 5.4 | 450 | 6 | AAI49541 | Tet opera |
| C 40 | 310.4 | 5.4 | 450 | 8 | ACA94736 | Minimal C |
| C 41 | 310.4 | 5.4 | 450 | 8 | ABX15339 | Human cyt |
| C 42 | 310.4 | 5.4 | 450 | 8 | ABX16489 | CMV promo |
| C 43 | 310.4 | 5.4 | 450 | 13 | ADR88197 | PhCMV*-2, |
| C 44 | 309.4 | 5.4 | 447 | 4 | AAD09962 | Tet respo |
| C 45 | 309.4 | 5.4 | 450 | 2 | AAT06869 | PhCMV*-1 |

ALIGNMENTS

RESULT 1

ADI28886
ID ADI28886 standard; DNA; 5735 BP.
XX
AC ADI28886;
XX
DT 22-APR-2004 (first entry)
XX
DE Inducible, cardiac-preferred promoter MHCminTetO.
XX
KW MHCminTetO; promoter; mouse; myosin; cardiomyopathy; cardiovascular-gen.;
KW cardiac; antiangiinal; gene therapy; transgenic; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
repeat_region 4282..4574
FT /*tag= b.
FT /rpt_family= "DIRECT"
FT /note= "TetO binding site"
FT repeat_unit 4282..4323
FT /*tag= a
XX
PN WO2004005474-A2.
XX
PD 15-JAN-2004.
XX
PF 03-JUL-2003; 2003WO-US021035.
XX
PR 03-JUL-2002; 2002US-0393525P.
XX
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI Robbins J;
XX
DR WPI; 2004-091352/09.
XX
PT New isolated nucleic acid molecule having inducible cardiac-preferred
PT expression, useful for diagnosing or treating cardiac diseases, including
PT ischemic heart disease, angina pectoris, myocardial infarction and
XX endocarditis.
PS Claim 1; SEQ ID NO 1; 69pp; English.

xx The present sequence is that of MHCminTetO, an inducible, cardiac-
cc preferred promoter sequence derived from the mouse alpha-myosin heavy
cc chain promoter sequence. The native sequence was modified by insertion of
cc a 7-repeat TetO binding site sequence. The promoter comprises a responder
cc locus that is a copy number dependent, position independent locus in
cc which various transgenes can be inserted. When uninduced, these
cc transgenes are silent. When induced, the transgenes are very active.
cc These genes can then be turned off using the inducible system. The
cc promoter is useful for expressing operably linked sequences in a cardiac
cc tissue-preferred expression pattern. Expression cassettes, host cells and
cc transgenic animals are provided. The transgenic animals exhibit inducible
cc cardiac-preferred expression of a nucleotide sequence of interest, e.g.
cc ELG1a or glycogen synthase kinase 3-beta. These animals may have an
cc altered susceptibility to cardiopathology and may be useful for
cc identifying anti-cardiopathic compounds. The cardiopathology is
cc especially a cardiomyopathy such as familial hypertrophic
cc cardiomyopathies, dilated cardiomyopathies, peripartum cardiomyopathy,
cc restrictive cardiomyopathies, ischaemic heart disease, angina pectoris,
cc myocardial infarction, hypertensive heart disease and endocarditis
cc (claimed).

xx
SQ Sequence 5735 BP; 1458 A; 1455 C; 1545 G; 1277 T; 0 U; 0 Other;

Query Match 100.0%; Score 5735; DB 12; Length 5735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCTCGAAGTTCACACAGAGTCTCCACCACAGGTGCGCTCCTCAATTTTCAGT 60
Db 1 GGATCTCGAAGTTCACACAGAGTCTCCACCACAGGTGCGCTCCTCAATTTTCAGT 60

QY 61 TTCCATGCTTGTCTTCAACATGCTGGCTCCCGACAGCTAATTTGGACTTTGTTTTAT 120
Db 61 TTCCATGCTTGTCTTCAACATGCTGGCTCCCGACAGCTAATTTGGACTTTGTTTTAT 120

QY 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGTGTACCCAGCTCTAAGGGTGCCCGTGA 180
Db 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGTGTACCCAGCTCTAAGGGTGCCCGTGA 180

QY 181 AGCCCTCAGACTTGAGGCTTTGCAACAGCCCTTTAGTGGAGCAGAAATAAGCAATTT 240
Db 181 AGCCCTCAGACTTGAGGCTTTGCAACAGCCCTTTAGTGGAGCAGAAATAAGCAATTT 240

QY 241 TCCTTAAAGCAAAATCTGCTCTAGACTCTTCTTCTGACCTCGCTCGCTGGGCTCT 300
Db 241 TCCTTAAAGCAAAATCTGCTCTAGACTCTTCTTCTGACCTCGCTCGCTGGGCTCT 300

QY 301 AGGGTGGGAGGTGGGGCTTGGAAGAGAGAGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
Db 301 AGGGTGGGAGGTGGGGCTTGGAAGAGAGAGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360

QY 361 GGCCCTGTGAAGTTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420
Db 361 GGCCCTGTGAAGTTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420

QY 421 AAACATACGAAGTGTACTCTCTTGTGACTTCCCGAGCCGACGATCTGTCAGTT 480
Db 421 AAACATACGAAGTGTACTCTCTTGTGACTTCCCGAGCCGACGATCTGTCAGTT 480

QY 481 GAAACAGGATTTAGAGAGCCCTTGAACTCAGCTGAACTCTGAAGCTCATCCACCAAGCA 540
Db 481 GAAACAGGATTTAGAGAGCCCTTGAACTCAGCTGAACTCTGAAGCTCATCCACCAAGCA 540

QY 541 AGCAGCTAGGTGCGCATCTGCTAGTTAGTATCTTACGCTGATATATGACAGCTGGGCCAC 600
Db 541 AGCAGCTAGGTGCGCATCTGCTAGTTAGTATCTTACGCTGATATATGACAGCTGGGCCAC 600

QY 601 AGAAGTCTGGGGTGTAGGAACTGACCTGACTTTTTCAGTCCGCAAGGATGACCCCC 660
Db 601 AGAAGTCTGGGGTGTAGGAACTGACCTGACTTTTTCAGTCCGCAAGGATGACCCCC 660

QY 661 TCAGCAGATGTAGTAATGTCCCTTTAGATCCCATCCAGGAGGTCTCTTAAGAGGACATG 720

Db 661 TCAGCAGATGTAGTAATGTCCCTTTAGATCCCATCCAGGAGGTCTCTTAAGAGGACATG 720

QY 721 GGATGAGAGATGTAGTCAATGTGGCATTTCCAAACACAGCTATCCACAGGTGCCCTTGGCCC 780
Db 721 GGATGAGAGATGTAGTCAATGTGGCATTTCCAAACACAGCTATCCACAGGTGCCCTTGGCCC 780

QY 781 TTCCACTTAGCCAGGAGACAGTAACTTACCTTATCTTTCTCTCCCTCCCATCTCCCCAG 840
Db 781 TTCCACTTAGCCAGGAGACAGTAACTTACCTTATCTTTCTCTCCCTCCCATCTCCCCAG 840

QY 841 GACACACCCCTGTGTCAGTATTCATTTCTTCTTCACTCCCTCTGACTTCCAT 900
Db 841 GACACACCCCTGTGTCAGTATTCATTTCTTCTTCACTCCCTCTGACTTCCAT 900

QY 901 TTGCAAGGCTTTTACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTTGGTGGGAA 960
Db 901 TTGCAAGGCTTTTACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTTGGTGGGAA 960

QY 961 GCCATCTCAAGAGAAAAGCAGACAAACAGGGGACAGATTTTGGAAAGGATCAGGAACATAA 1020
Db 961 GCCATCTCAAGAGAAAAGCAGACAAACAGGGGACAGATTTTGGAAAGGATCAGGAACATAA 1020

QY 1021 TCACCTGCGGGCTTGGGGGTAGAAAAAGAGTGAAGTGAAGTCCGCTTCCAGCTTAAGCACAAGC 1080
Db 1021 TCACCTGCGGGCTTGGGGGTAGAAAAAGAGTGAAGTGAAGTCCGCTTCCAGCTTAAGCACAAGC 1080

QY 1081 TAGTCCCAGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
Db 1081 TAGTCCCAGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140

QY 1141 TGAAGAACAATGGGATTTGGAAGACATCTCTTTTGTAGTCTCCCTCAAACCCACCTACAGAC 1200
Db 1141 TGAAGAACAATGGGATTTGGAAGACATCTCTTTTGTAGTCTCCCTCAAACCCACCTACAGAC 1200

QY 1201 ACACCTGCTGTGTGCCAGACTCTTGTCAACAGCCCTCTGTGTCTGTGACCTTCCAGCTAG 1260
Db 1201 ACACCTGCTGTGTGCCAGACTCTTGTCAACAGCCCTCTGTGTCTGTGACCTTCCAGCTAG 1260

QY 1261 GCAACAGAGATGGGCTGCTGTGAGGATGAAGAGTGTGTTTACCANTAGCAAAAACAG 1320
Db 1261 GCAACAGAGATGGGCTGCTGTGAGGATGAAGAGTGTGTTTACCANTAGCAAAAACAG 1320

QY 1321 CAGGGGAGGAGAACAGAGAACGAAATAAGGAAGAGAGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 CAGGGGAGGAGAACAGAGAACGAAATAAGGAAGAGAGAGGAGGAGGAGGAGGAGGAG 1380

QY 1381 TGCAAGTCAAGAGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381 TGCAAGTCAAGAGAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

QY 1441 AGATTTGGGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 1441 AGATTTGGGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

QY 1501 ATACCTCATCTCCCTCTCCCAATTAAGCCCACTCTTCTTCTAGATCAGACTGAGCTGC 1560
Db 1501 ATACCTCATCTCCCTCTCCCAATTAAGCCCACTCTTCTTCTAGATCAGACTGAGCTGC 1560

QY 1561 AGCGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 1561 AGCGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

QY 1621 GCAACCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db 1621 GCAACCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680

QY 1681 CTGTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 CTGTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740

QY 1741 ATGGCGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1741 ATGGCGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

[illegible][illegible]

RESULT 2

RESOLU 2
AAD02408

ID AAD02408 standard; DNA; 5443 BP.

AC AAD02408;

DT 24-APR-2001 (first entry)

DE Mouse alpha-cardiac myosin heavy chain (MHC) promoter.

Mouse: cvcln D2: CYCD2: cardiomyocyte cell: screening

infarct: cardiomyopathy; alpha-cardiac myosin heavy chain; MHC; ds.

OS Mus musculus.

PN WO200078119-A

PD 28-DEC-2000.

19-JUN-2000;

PR 18-JUN-1999; 99US-0139942P.

PA (ADRE-) ADVANCED RES & TECHNOC

XX Field LJ, Paeumarthi KBS;
XX WPI; 2001-102575/11.
DR Increasing proliferative potential of cardiomyocyte cell which is used
XX for screening activity of biological or pharmacological agent, involves
PT increasing the level of cyclin D2 activity in cardiomyocyte cell.
PT
XX
PS Example 1; Page 62-66; 67pp; English.
XX
CC The present sequence is mouse alpha-cardiac myosin heavy chain (MHC)
CC promoter. This sequence is used in the preparation of a MHC-CYC2 fusion
CC gene. The cyclin D2 (CYC2) cDNA sequence is used to increase the
CC proliferative potential of a cardiomyocyte cell by increasing the level
CC of cyclin D2 activity in the cardiomyocyte cell. The transgenic animals
CC expressing cyclin D2 have sustained atrial and ventricular cardiomyocyte
CC DNA synthesis. Cardiomyocyte cells with enhanced proliferative potential
CC are useful for screening the activity of biological or pharmacological
CC agents on cardiomyocyte cells. Genetically modified cardiomyocyte cells
CC are useful for delivering therapeutics to mammals. The cells are also
CC useful to target an improvement of the contractile function of the heart
CC of the patient, for e.g. in the treatment of contractile losses due to
CC infarcts or cardiomyopathies
XX
SQ Sequence 5443 BP; 1352 A; 1399 C; 1492 G; 1200 T; 0 U; 0 Other;
Query Match 88.0%; Score 5045; DB 5; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;
QY 1 GGATCTCCAGGTCACACAGGGTCTCCACCACAGGTGCGCTAGTCTCAATTTTCACT 60
DB 1 GGATCTCCAGGTCACACAGGGTCTCCACCACAGGTGCGCTAGTCTCAATTTTCACT 60
QY 61 TTCCATGCTTGTCTCACAATGCTGGCTCCCCAGAGCTAATTTGGACTTTGTTTAT 120
DB 61 TTCCATGCTTGTCTCACAATGCTGGCTCCCCAGAGCTAATTTGGACTTTGTTTAT 120
QY 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGTGTACCCAGCTCTAAGGGTGCCGTGA 180
DB 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGTGTACCCAGCTCTAAGGGTGCCGTGA 180
QY 181 AGCCCTCAGACTGGAGCTTTGCAACAGCCCTTTAGTGGAGCAGATAAAGCAATTT 240
DB 181 AGCCCTCAGACTGGAGCTTTGCAACAGCCCTTTAGTGGAGCAGATAAAGCAATTT 240
QY 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGGTCCCTGGGCTCT 300
DB 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGGTCCCTGGGCTCT 300
QY 301 AGGGTGGGGAGGTGGGGCTTGGAAAGAGAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
DB 301 AGGGTGGGGAGGTGGGGCTTGGAAAGAGAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
QY 361 GGCCCTGTGAATTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420
DB 361 GGCCCTGTGAATTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420
QY 421 AAACATACAGAGTGTACTCTCTTGTGACTTCCCGAGCCAGTACCTGTCAAGTT 480
DB 421 AAACATACAGAGTGTACTCTCTTGTGACTTCCCGAGCCAGTACCTGTCAAGTT 480
QY 481 GAAACAGGATTTAGAGAGCCCTCTGAACCTCACTGAACTCTGAAGCTCATCCACCAAGCA 540
DB 481 GAAACAGGATTTAGAGAGCCCTCTGAACCTCACTGAACTCTGAAGCTCATCCACCAAGCA 540
QY 541 AGCACCCTAGGTGCCACTGCTAGTATGATATCTTACCTGATATATATGAGAGCTGGGCCAC 600
DB 541 AGCACCCTAGGTGCCACTGCTAGTATGATATCTTACCTGATATATATGAGAGCTGGGCCAC 600
QY 601 AGAGTCTGGGTGTAGGACTGACAGTACTTTTTCAGTCGGCAAGGTATGACCCCC 660
DB 601 AGAGTCTGGGTGTAGGACTGACAGTACTTTTTCAGTCGGCAAGGTATGACCCCC 660
DB 601 AGAGTCTGGGTGTAGGACTGACAGTACTTTTTCAGTCGGCAAGGTATGACCCCC 660
QY 661 TCAGCAGATGTAGTAATGTCCCTTAGATCCCATCCAGGAGGTCTCTAAGAGGACATG 720
DB 661 TCAGCAGATGTAGTAATGTCCCTTAGATCCCATCCAGGAGGTCTCTAAGAGGACATG 720
QY 721 GGATGAGAGATGTAGTCAATGTGGCAATCCAAACACAGCTATCCACAGTGTCTCTGCCCC 780
DB 721 GGATGAGAGATGTAGTCAATGTGGCAATCCAAACACAGCTATCCACAGTGTCTCTGCCCC 780
QY 781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCTATCTTTCTCTCTCCCATCTCTCCAG 840
DB 781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCTATCTTTCTCTCTCCCATCTCTCCAG 840
QY 841 GACACACCCCTGTCTGCAGTATTCATTTCTTCCCTTCACTGCTCCCTCTGTGACTTCCAT 900
DB 841 GACACACCCCTGTCTGCAGTATTCATTTCTTCCCTTCACTGCTCCCTCTGTGACTTCCAT 900
QY 901 TTGCAAGGCTTTTGACCTCTGCACTCTGCAAGTGTGGAAGATAGATTTTGGAGGATCAGAA 960
DB 901 TTGCAAGGCTTTTGACCTCTGCACTCTGCAAGTGTGGAAGATAGATTTTGGAGGATCAGAA 960
QY 961 GCCATCTCAAGAGAAAGCAGACAACAGGGGGAACAGATTTTGGAAAGGATCAGAACTAAA 1020
DB 961 GCCATCTCAAGAGAAAGCAGACAACAGGGGGAACAGATTTTGGAAAGGATCAGAACTAAA 1020
QY 1021 TCATCGGGGGCTGGGGGTAGAAAAGAGAGTGAAGTCCGCTCCAGCTAAGCAAGC 1080
DB 1021 TCATCGGGGGCTGGGGGTAGAAAAGAGAGTGAAGTCCGCTCCAGCTAAGCAAGC 1080
QY 1081 TAGTCCCGAGACTCTGCCACAGCTGGCTGCTGGGGTAGCTTTAGGAATGTGGTCT 1140
DB 1081 TAGTCCCGAGACTCTGCCACAGCTGGCTGCTGGGGTAGCTTTAGGAATGTGGTCT 1140
QY 1141 TGAAGACAATGGGATTTGGAAGACATCTCTTTGAGTCTCTCCCTCAACCCACCTACAGAC 1200
DB 1141 TGAAGACAATGGGATTTGGAAGACATCTCTTTGAGTCTCTCCCTCAACCCACCTACAGAC 1200
QY 1201 ACATCTGTGTGGCAGACTCTCTTCAACAGCCCTCTGTGTTCTGACCTACCTAGCTAG 1260
DB 1201 ACATCTGTGTGGCAGACTCTCTTCAACAGCCCTCTGTGTTCTGACCTACCTAGCTAG 1260
QY 1261 GCACCCAGAGATGGGGCTGTGCTGAGGATGAAGTGGTTTCAATAGCAAAACAG 1320
DB 1261 GCACCCAGAGATGGGGCTGTGCTGAGGATGAAGTGGTTTCAATAGCAAAACAG 1320
QY 1321 CAGGGGAGAGAACAGAGAACGAAATAGGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1321 CAGGGGAGAGAACAGAGAACGAAATAGGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 1381 TGCACTCAGAGAGATGGGAGCCCAACACAGCTTTGAGCAGAGAGAAACAGAAAGGGAG 1440
DB 1381 TGCACTCAGAGAGATGGGAGCCCAACACAGCTTTGAGCAGAGAGAAACAGAAAGGGAG 1440
QY 1441 AGATTTCTGGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 AGATTTCTGGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 ATACCTCTATCCCGTCTCCCAATTAAGCCCACTCTTCTCTAGATCAGACCTGAGCTGC 1560
DB 1501 ATACCTCTATCCCGTCTCCCAATTAAGCCCACTCTTCTCTAGATCAGACCTGAGCTGC 1560
QY 1561 AGCGAAGAGACCCGTAGGAGGAGTCACTGGATGAAGGAGATGTGTGAGAGAGTCCAGG 1620
DB 1561 AGCGAAGAGACCCGTAGGAGGAGTCACTGGATGAAGGAGATGTGTGAGAGAGTCCAGG 1620
QY 1621 GCAACCTTAAGAGCCAGAGCTTAAGAGCAGAGATTAAGTGTCTCAAGGTGGCCAGG 1680
DB 1621 GCAACCTTAAGAGCCAGAGCTTAAGAGCAGAGATTAAGTGTCTCAAGGTGGCCAGG 1680
QY 1681 CTGTGCACACAGAGGGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 1740
DB 1681 CTGTGCACACAGAGGGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 1740

QY 1741 ATGGCGGGGGGGGATTTCTGGGGGGCGGAGAGAGAGTGAAGAGAGCTTGGAAACAG 1800
DB 1741 ATGGCGGGGGGGGGAATTTCTGGGGGGGGGAGAGAGAGAGTGAAGAGAGCTTGGAAACAG 1800
QY 1801 AGAATCTGGAAGCGCTGGAACCGATACATCAATAAAGGGAAGAAACCCAGGCTACCTTTAGATG 1860
DB 1801 AGAATCTGGAAGCGCTGGAACCGATACATCAATAAAGGGAAGAAACCCAGGCTACCTTTAGATG 1860
QY 1861 TAAATCATGAAGACAGGGAGAAAGGAAGCTGGAGAGAGTAGAAGGACCCCGGGGCGAGA 1920
DB 1861 TAAATCATGAAGACAGGGAGAAAGGAAGCTGGAGAGAGTAGAAGGACCCCGGGGCGAGA 1920
QY 1921 CATGGAAGCAAGGCAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTCTGAAAGCAGAG 1980
DB 1921 CATGGAAGCAAGGCAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTCTGAAAGCAGAG 1980
QY 1981 CCCTGGTATGACACACAGAAACAGCAGAGGCTAGGGTTAATGTCGAGACAGGGAAACAGAAAG 2040
DB 1981 CCCTGGTATGACACACAGAAACAGCAGAGGCTAGGGTTAATGTCGAGACAGGGAAACAGAAAG 2040
QY 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAACAAAGGAATGGTCTCTTCAC 2100
DB 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAACAAAGGAATGGTCTCTTCAC 2100
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACCTAGAAATGTTTCATCAGACTGCGAGGC 2160
DB 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACCTAGAAATGTTTCATCAGACTGCGAGGC 2160
QY 2161 TGGCTTGGGAGGCGAGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAGGGGGCTTAGGA 2220
DB 2161 TGGCTTGGGAGGCGAGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAGGGGGCTTAGGA 2220
QY 2221 AAGGAAGAGGGGAAACAGAGCCACACAGAGGGGAGAGCCAGAACTGAGTTAACTC 2280
DB 2221 AAGGAAGAGGGGAAACAGAGCCACACAGAGGGGAGAGCCAGAACTGAGTTAACTC 2280
QY 2281 CTTCCTCTGTGATCTCCATAGGAGGAGTGGGAACCTGTGACCAACATCCCCCATGA 2340
DB 2281 CTTCCTCTGTGATCTCCATAGGAGGAGTGGGAACCTGTGACCAACATCCCCCATGA 2340
QY 2341 GCCCCACTACCCATACCAAGTTTGGCCCTGAGTGGCATTCTAGTTCCCTGAGGACAGAG 2400
DB 2341 GCCCCACTACCCATACCAAGTTTGGCCCTGAGTGGCATTCTAGTTCCCTGAGGACAGAG 2400
QY 2401 CCTGGCCCTTTGTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTGAAT 2460
DB 2401 CCTGGCCCTTTGTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTATCAT 2460
QY 2461 GCCCTCAAGAGCTTGAGAACAGGACAGTGCATATTTAGGCCATGGGCTAACCCCTGGAGCT 2520
DB 2461 GCCCTCAAGAGCTTGAGAACAGGACAGTGCATATTTAGGCCATGGGCTAACCCCTGGAGCT 2520
QY 2521 TGCAACAGGAGCCCTCAAGTGAACCTCCAGGGACACAGCTGACAGAGTGCGCTTTATCC 2580
DB 2521 TGCAACAGGAGCCCTCAAGTGAACCTCCAGGGACACAGCTGACAGAGTGCGCTTTATCC 2580
QY 2581 CCAAAGACCAACCATTTGGCATAGTGGCTGCAAAATGGGAATGCAAGTTGAAATCAGGTC 2640
DB 2581 CCAAAGACCAACCATTTGGCATAGTGGCTGCAAAATGGGAATGCAAGTTGAAATCAGGTC 2640
QY 2641 CCTTCAAGAACTGTCATGCAAGACCTTAAGACCCCTGGAGAGGGGTATGCTCCTGCCC 2700
DB 2641 CCTTCAAGAACTGTCATGCAAGACCTTAAGACCCCTGGAGAGGGGTATGCTCCTGCCC 2700
QY 2701 CCACCACCAATAAGGGGAGTGAACCTATCTAGGGGGCTGGGACCTTGGGGAGACACCCAC 2760
DB 2701 CCACCACCAATAAGGGGAGTGAACCTATCTAGGGGGCTGGGACCTTGGGGAGACACCCAC 2760
QY 2761 ATTACTGAGGTGCTGAGCCCAAGAAAACCTGACCGGCCCTGTGCTGCGCCACCTCCACAC 2820
DB 2761 ATTACTGAGGTGCTGAGCCCAAGAAAACCTGACCGGCCCTGTGCTGCGCCACCTCCACAC 2820

QY 2821 TCTAGAGCTATATTGAGAGGTGA CAGTAGATAGGGTGGGAGCTGCTAGCAGGGAGAGTGT 2880
DB 2821 TCTAGAGCTATATTGAGAGGTGA CAGTAGATAGGGTGGGAGCTGCTAGCAGGGAGAGTGT 2880
QY 2881 TCCTGGGTGAGAGGGTGTAGGGAAAGCCAGAGACAGGGGAGTCTGGCTTTGTCTTCCTGAA 2940
DB 2881 TCCTGGGTGAGAGGGTGTAGGGAAAGCCAGAGACAGGGGAGTCTGGCTTTGTCTTCCTGAA 2940
QY 2941 CACAATCTCTACTTACTTAAACAGGATGACCTGCTAAAGACCAACATCTACGACCTC 3000
DB 2941 CACAATCTCTACTTACTTAAACAGGATGACCTGCTAAAGACCAACATCTACGACCTC 3000
QY 3001 TGAAGAAGACAGACCCCTGGAGGACAGGGTGTCTCTGAGCCTTGGGTGCTTGATGGTG 3060
DB 3001 TGAAGAAGACAGACCCCTGGAGGACAGGGTGTCTCTGAGCCTTGGGTGCTTGATGGTG 3060
QY 3061 CCAAAAGGAGGGCATGAGTGTGAGTAAAGGCCCAAGGCGGTAGAGAAAGGGCAGCTTG 3120
DB 3061 CCAAAAGGAGGGCATGAGTGTGAGTAAAGGCCCAAGGCGGTAGAGAAAGGGCAGCTTG 3120
QY 3121 GGAAGGGGTGAGTCTGAGAGCCCTATCCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
DB 3121 GGAAGGGGTGAGTCTGAGAGCCCTATCCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
QY 3181 TAAATCTCTGGGCTGCGCAGGCAATTCAGGACAGCACCTGCATCTCTGGCAGCCTGGGGA 3240
DB 3181 TAAATCTCTGGGCTGCGCAGGCAATTCAGGACAGCACCTGCATCTCTGGCAGCCTGGGGA 3240
QY 3241 GGCGGAAGGAGCAACCCGCCACTTATACCTTTCTCCCTCAGGCCCAAGGATTAACACT 3300
DB 3241 GGCGGAAGGAGCAACCCGCCACTTATACCTTTCTCCCTCAGGCCCAAGGATTAACACT 3300
QY 3301 CTGGCCTTCCCTTCCCTCCCTCAGGAGTGGAGGGTTCAGAGGGAGGGTAAAAA 3360
DB 3301 CTGGCCTTCCCTTCCCTCCCTCAGGAGTGGAGGGTTCAGAGGGAGGGTAAAAA 3360
QY 3361 CCTACATGTCAAAACATCATGTGTCAGATATATGGAATCTGATGATGATGATGATGATGATG 3420
DB 3361 CCTACATGTCCAAACATCATGTGTCAGATATATGGAATCTGATGATGATGATGATGATG 3420
QY 3421 GGAATCTGCAAGCTTAACTGGGTTAATGTGTAAGTCTGTGTGATGTGTGTGTGTGTGTGTG 3480
DB 3421 GGAATCTGCAAGCTTAACTGGGTTAATGTGTAAGTCTGTGTGATGTGTGTGTGTGTGTGTG 3480
QY 3481 ACTGAARAGGGCATGCTGTCAGCTGTCAGTTCCTGCGTGAGGTACCAAGACTGCA 3540
DB 3481 ACTGAARAGGGCATGCTGTCAGCTGTCAGTTCCTGCGTGAGGTACCAAGACTGCA 3540
QY 3541 GGTTTGTGTGTAATAATGCCCAGGCAAGTGGGTGAATCCCTTCATGGTTTAAAGAGAT 3600
DB 3541 GGTTTGTGTGTAATAATGCCCAGGCAAGTGGGTGAATCCCTTCATGGTTTAAAGAGAT 3600
QY 3601 TGGATGATGGCTGCATCTCAAGGACCAATGGAATAAGAAATGGAACACTCTATATGTGTCT 3660
DB 3601 TGGATGATGGCTGCATCTCAAGGACCAATGGAATAAGAAATGGAACACTCTATATGTGTCT 3660
QY 3661 CTAAAGCTAAGGTAGCAAGGCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
DB 3661 CTAAAGCTAAGGTAGCAAGGCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
QY 3721 TACAGACAGTATCTGTACAGAGTAAGGAGAGAGGGGGGTGTAGAAATTTCTCTTACTA 3780
DB 3721 TACAGACAGTATCTGTACAGAGTAAGGAGAGAGGGGGGTGTAGAAATTTCTCTTACTA 3780
QY 3781 TCAAGGAGGAACTGAGTGTGACCTGCAAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840
DB 3781 TCAAGGAGGAACTGAGTGTGACCTGCAAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840
QY 3841 TGTCTCTGGGAGCCAGCTGTCAGGCTTTCAGGCTTTCAGGAGTGTAGGAGGCTCCCTCA 3900
DB 3841 TGTCTCTGGGAGCCAGCTGTCAGGCTTTCAGGCTTTCAGGAGTGTAGGAGGCTCCCTCA 3900
QY 3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAAGGAGAGCTTGGATGGGAGC 3960

Db 1681 CTGTGCACACAGAGGGTTCGAGACTGGTGGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
Qy 1741 ATGGCGGGGGGGGGGATTTTGGGGGGGGGAGAGAGAAAGGTGAGAAAGCCTGGAACAG 1800
Db 1741 ATGGCGGGGGGGGGGATTTTGGGGGGGGGAGAGAGAAAGGTGAGAAAGCCTGGAACAG 1800
Qy 1801 AGAATCTGGAAGCGCTGGAACGATACCATAAAGGGAAGAACCCAGGCTACCTTTAGATG 1860
Db 1801 AGAATCTGGAAGCGCTGGAACGATACCATAAAGGGAAGAACCCAGGCTACCTTTAGATG 1860
Qy 1861 TAAATCATGAAGACAGGAGAGAAAGCTGGAGAGAGTGAAGGACCCCGGGCAAGA 1920
Db 1861 TAAATCATGAAGACAGGAGAGAAAGCTGGAGAGAGTGAAGGACCCCGGGCAAGA 1920
Qy 1921 CATGGAACAGGACAAAGCCAGGTTGAGGCTCCGTGAAATCAGCCTGCTGAAGCAGAG 1980
Db 1921 CATGGAACAGGACAAAGCCAGGTTGAGGCTCCGTGAAATCAGCCTGCTGAAGCAGAG 1980
Qy 1981 CCCTGGTATGAGCACCAAGACAGCAGAGGCTAGGGTTAATGTCGAGACAGAGGAACAGAG 2040
Db 1981 CCCTGGTATGAGCACCAAGACAGCAGAGGCTAGGGTTAATGTCGAGACAGAGGAACAGAG 2040
Qy 2041 GTAGACACAGGAACACAGAGACAGGAGGAGCCAGGTAACAAAGGAATGGTCTTCTCAC 2100
Db 2041 GTAGACACAGGAACACAGAGACAGGAGGAGCCAGGTAACAAAGGAATGGTCTTCTCAC 2100
Qy 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACCTAGAAATGTTTCATCAGACTGAGGGC 2160
Db 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACCTAGAAATGTTTCATCAGACTGAGGGC 2160
Qy 2161 TGGCTTGGGAGGCGAGCTGGAAGAGTATGTAGAGCCAGGGGAGAACAGGGGGCTTAGGA 2220
Db 2161 TGGCTTGGGAGGCGAGCTGGAAGAGTATGTAGAGCCAGGGGAGAACAGGGGGCTTAGGA 2220
Qy 2221 AAGGAAGAGGGGCAAAACAGGCCACACAAAGAGGCGAGAGCCAGAACTGAGTTAACTC 2280
Db 2221 AAGGAAGAGGGGCAAAACAGGCCACACAAAGAGGCGAGAGCCAGAACTGAGTTAACTC 2280
Qy 2281 CTTCCCTTGTGATCTTCATAGAGGAGGAGTGGGAACTCTGTGACCAACATCCCCCATGA 2340
Db 2281 CTTCCCTTGTGATCTTCATAGAGGAGGAGTGGGAACTCTGTGACCAACATCCCCCATGA 2340
Qy 2341 GCGCCACTACCCATACCAAGTTTGGCTGAGTGGCAATCTAGTTCCCTGAGGACAGAG 2400
Db 2341 GCGCCACTACCCATACCAAGTTTGGCTGAGTGGCAATCTAGTTCCCTGAGGACAGAG 2400
Qy 2401 CCTGGCTTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTGAAT 2460
Db 2401 CCTGGCTTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTGAAT 2460
Qy 2461 GCGCTCAAGAGCTTGAGAACGAGGAGTGACATATTAGGCCATGGGCTAAACCTGGAGCT 2520
Db 2461 GCGCTCAAGAGCTTGAGAACGAGGAGTGACATATTAGGCCATGGGCTAAACCTGGAGCT 2520
Qy 2521 TGCCACAGGAGCCTCAAGTACCTCCAGGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
Db 2521 TGCCACAGGAGCCTCAAGTACCTCCAGGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
Qy 2581 CCAAGAGCAACCATTTGGCATAGTGGCTGCAATGGGAATGCAAGGTGGAATCAGGTC 2640
Db 2581 CCAAGAGCAACCATTTGGCATAGTGGCTGCAATGGGAATGCAAGGTGGAATCAGGTC 2640
Qy 2641 CTTTCAAGAAATATGTCATGCAAGACTTAAGACCCCTGGAGAGAGGGGTATGCTCTGCC 2700
Db 2641 CTTTCAAGAAATATGTCATGCAAGACTTAAGACCCCTGGAGAGAGGGGTATGCTCTGCC 2700
Qy 2701 CCACCCACCATAGGGAGTGAATATCTTAGGGGCTGGGACCTTGGGAGAGACACAC 2760
Db 2701 CCACCCACCATAGGGAGTGAATATCTTAGGGGCTGGGACCTTGGGAGAGACACAC 2760
Qy 2761 ATTACTGAGTGTGAGCCAGAAAAACTGACCGCCCTGTGTCTGCTGCCACCTCCAC 2820
Db 2761 ATTACTGAGTGTGAGCCAGAAAAACTGACCGCCCTGTGTCTGCTGCCACCTCCAC 2820

Qy 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGGAGCTGCTAGCAGGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGGAGCTGCTAGCAGGAGAGTGT 2880
Qy 2881 TCTTGGGTGTGAGGGTGTAGGGGAAGCCAGACAGGAGGAGTCTGGCTTTGTCTCTGAA 2940
Db 2881 TCTTGGGTGTGAGGGTGTAGGGGAAGCCAGACAGGAGGAGTCTGGCTTTGTCTCTGAA 2940
Qy 2941 CACAATGTCTACTTAGTTTATACAGGCAATGCTGCTAAAGACCCAAATCTACGACCTC 3000
Db 2941 CACAATGTCTACTTAGTTTATACAGGCAATGCTGCTAAAGACCCAAATCTACGACCTC 3000
Qy 3001 TGAAGAGACAGACGCCCTGGAGGACAGGGTGTCTCTGAGCCTTGGGCTTGTGATGTTG 3060
Db 3001 TGAAGAGACAGACGCCCTGGAGGACAGGGTGTCTCTGAGCCTTGGGCTTGTGATGTTG 3060
Qy 3061 CCACAAAGAGGGGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCACTTG 3120
Db 3061 CCACAAAGAGGGGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCACTTG 3120
Qy 3121 GGAAGGGGTGAGTCTGCAAGAGCCCTATCCATGGAATCTGAGCCTGGGCACTGTTG 3180
Db 3121 GGAAGGGGTGAGTCTGCAAGAGCCCTATCCATGGAATCTGAGCCTGGGCACTGTTG 3180
Qy 3181 TAAATCTCTGGGCTGCCAGGCAATCAAGCAGACACCTGCATCTCTGGCAGCCTGGGA 3240
Db 3181 TAAATCTCTGGGCTGCCAGGCAATCAAGCAGACACCTGCATCTCTGGCAGCCTGGGA 3240
Qy 3241 GGCAGAGGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCAGGATTAACACT 3300
Db 3241 GGCAGAGGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCAGGATTAACACT 3300
Qy 3301 CTGGCTTCCCTTCCACCTCCCATCAGAGTGGAGGGTTGCAGAGGAGGGTAAAAA 3360
Db 3301 CTGGCTTCCCTTCCACCTCCCATCAGAGTGGAGGGTTGCAGAGGAGGGTAAAAA 3360
Qy 3361 CCTACATGTCACAAATCATCATGTCGACGATATATGGAATCAGTATGTGTAGAGGCAAGAA 3420
Db 3361 CCTACATGTCACAAATCATCATGTCGACGATATATGGAATCAGTATGTGTAGAGGCAAGAA 3420
Qy 3421 GGAATCTGCAAGGCTTAACTGGGTTAATGTTAAAGTCTGTGTCATCTGTGTGTCTG 3480
Db 3421 GGAATCTGCAAGGCTTAACTGGGTTAATGTTAAAGTCTGTGTCATCTGTGTGTCTG 3480
Qy 3481 ACTGAAAACGGGCTGCTGTCAGCTGTTCAAGTCTGTGTCGAGGTTACCAGACTGCA 3540
Db 3481 ACTGAAAACGGGCTGCTGTCAGCTGTTCAAGTCTGTGTCGAGGTTACCAGACTGCA 3540
Qy 3541 GGTGTTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
Db 3541 GGTGTTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
Qy 3601 TGGATGATGGCTCCATCTCAAGGACCAATGGAATAGAAATAGAAATGGAACCTTATGCT 3660
Db 3601 TGGATGATGGCTCCATCTCAAGGACCAATGGAATAGAAATAGAAATGGAACCTTATGCT 3660
Qy 3661 CTAAGCTAAGGTAGCAGGCTTTTGGAGACACCTGTCTAGAGATGTTGGGCAACAGAGAC 3720
Db 3661 CTAAGCTAAGGTAGCAGGCTTTTGGAGACACCTGTCTAGAGATGTTGGGCAACAGAGAC 3720
Qy 3721 TACAGACAGTATCTGTACAGATTAAGGAGAGAGAGGGGTGTAGAAATCTCTTACTA 3780
Db 3721 TACAGACAGTATCTGTACAGATTAAGGAGAGAGAGGGGTGTAGAAATCTCTTACTA 3780
Qy 3781 TCAAGGGAACTCAGTGTGCACTGTCAGAACTGTCCTCCACACATCATGACTT 3840
Db 3781 TCAAGGGAACTCAGTGTGCACTGTCAGAACTGTCCTCCACACATCATGACTT 3840
Qy 3841 TGTCTCTGGGAGGCGAGCAGTGTGGAATCTCAGGCTCTGAGAGAGTAGGAGGCTCCCTCA 3900
Db 3841 TGTCTCTGGGAGGCGAGCAGTGTGGAATCTCAGGCTCTGAGAGAGTAGGAGGCTCCCTCA 3900

```
QY 3901 GCCTGAAGCTATGAGATAGCAGGTTGAAAGGGGGAAGGAGAGCCTGGATGGGAGC 3960
Db 3901 GCCTGAAGCTATGAGATAGCAGGTTGAAAGGGGGAAGGAGAGCCTGGATGGGAGC 3960
QY 3961 TTGTGTGTTGAGGAGGAGGAGCAGATATTAAGCCCTGGAAGAGAGGTGACCCCTTACCCAG 4020
Db 3961 TTGTGTGTTGAGGAGGAGGAGCAGATATTAAGCCCTGGAAGAGAGGTGACCCCTTACCCAG 4020
QY 4021 TTGTTCACCTCACCTTCAGATTAATAAATACTGAGGTAAGGGCTGGGTAGGGAGGTG 4080
Db 4021 TTGTTCACCTCACCTTCAGATTAATAAATACTGAGGTAAGGGCTGGGTAGGGAGGTG 4080
QY 4081 GTGTGAGACGCTCTCTCTCTCTGTCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGT 4140
Db 4081 GTGTGAGACGCTCTCTCTCTCTCTATCTGCCCCATCGGCCCTTTGGGGAGGAGGAATGT 4140
QY 4141 GCCAAGGACTAAATAAAGGCCATGAGCCAGAGAGGGCGAGGGCAACAGACTTTTCATGG 4200
Db 4141 GCCAAGGACTAAATAAAGGCCATGAGCCAGAGAGGGCGAGGGCAACAGACTTTTCATGG 4200
QY 4201 GCBAACCTTTGGGGCCCGTAGTGATCGATTGACAAGAACTCGCCCAATCGATACCCTTCTTC 4260
Db 4201 GCBAACCTTTGGGGCCCGTAGTGATCGATTGACAAGAACTCGCCCAATCGATACCCTTCTTC 4260
QY 4261 TTCTAAACGACAGAGGAGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGA 4320
Db 4261 GGAGCCAGACAGAGGAGGAGTGGAGGGAG----- 4291
QY 4321 AAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTC 4380
Db 4321 ----- 4291
QY 4381 CCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440
Db 4381 ----- 4291
QY 4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTT 4500
Db 4441 ----- 4291
QY 4501 CCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560
Db 4501 ----- 4291
QY 4561 AGAGAAAAGTGAAAGTCGAGTCGCTACGACGAGGAGCTCCAAATTTAGCAGCAGGCA 4620
Db 4561 ----- 4291
QY 4621 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTGAGAGATTTCTCCAAACCCA 4680
Db 4621 ----- 4291
QY 4681 GGTAGAGGAGTTTCGGGTGGGGCTCTTCAACCACACCGAGCTCTCCCACTAGAGAA 4740
Db 4681 GGTAGAGGAGTTTCGGGTGGGGCTCTTCAACCACACCGAGCTCTCCCACTAGAGAA 4740
QY 4741 GGAAACTGCCCTTCTCTGAAGTGGGTTTCAGCGCTGAGAGCTGTGAGAGGTGGGCTT 4800
Db 4741 GGAAACTGCCCTTCTCTGAAGTGGGTTTCAGCGCTGAGAGCTGTGAGAGGTGGGCTT 4800
QY 4801 CCACGAGCTGGGAGTTTCTCAGTGGCAGGAGGTTTCCACAAGAAACACTGGATGCCCT 4860
Db 4801 CCACGAGCTGGGAGTTTCTCAGTGGCAGGAGGTTTCCACAAGAAACACTGGATGCCCT 4860
QY 4861 TCCCTTACGCTGTCTTCTCTCAATCTTCTCTGGGAGTCTCTCCCGCTTTGGTTATC 4920
Db 4861 TCCCTTACGCTGTCTTCTCTCAATCTTCTCTGGGAGTCTCTCCCGCTTTGGTTATC 4920
QY 4921 TTGGCTCTTCTCTCAGCAGAGATTTGCCGTGTGCTGTCCACTCCACTCTTCTCTACTGT 4980
Db 4921 TTGGCTCTTCTCTCAGCAGAGATTTGCCGTGTGCTGTCCACTCCACTCTTCTCTACTGT 4980
QY 4981 CTCGCTGCTTGGCTTTCGCTTCTTTCGCTGTCTCTTCTTCCACCACTTCTCACTTACC 5040
```

```
Db 4689 CTCGCTGCTTGCCTTTCGCTTCTTTCGCTGTCTTCTTCTTCCACCACTTCTCACTTACC 4748
QY 5041 TTTTCTCCCTTCTCATTTGTAATTCATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5100
Db 4749 TTTTCTCCCTTCTCATTTGTAATTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4808
QY 5101 CTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5160
Db 4809 CTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4868
QY 5161 CTGTGTGAGAGTGTCTGAGAAATCACACCTGGGGTTCACACCTTATGTAAAAAATCTTCCA 5220
Db 4869 CTGTGTGAGAGTGTCTGAGAAATCACACCTGGGGTTCACACCTTATGTAAAAAATCTTCCA 4928
QY 5221 GTGAGCCACAGCTTCAGTGTCTGGGTGTCTCTTACCTTCTTCTTCTTCTTCTTCTTCTTCT 5280
Db 4929 GTGAGCCACAGCTTCAGTGTCTGGGTGTCTCTTACCTTCTTCTTCTTCTTCTTCTTCTTCT 4988
QY 5281 CTGTTCATCTGTCTGAGGATCTCTAGATTGGTCTTCCAGCTCTGTCTTCTTCTTCTTCTTCT 5340
Db 4989 CTGTTCATCTGTCTGAGGATCTCTAGATTGGTCTTCCAGCTCTGTCTTCTTCTTCTTCTTCT 5048
QY 5341 GCCTGTCTCTCTCTCTGTCAGCTGCGCACTGTGGTGCCTCTGTTCACAGCTGTGTCTTCTTCT 5400
Db 5049 GCCTGTCTCTCTCTCTGTCAGCTGCGCACTGTGGTGCCTCTGTTCACAGCTGTGTCTTCTTCT 5108
QY 5401 ATTCTTCAGGATCTCTGAAAAAGTTAACAGGTGAGAAATGTTTCCCTGTGAGACAGCAGA 5460
Db 5109 ATTCTTCAGGATCTCTGAAAAAGTTAACAGGTGAGAAATGTTTCCCTGTGAGACAGCAGA 5168
QY 5461 TCAGATTTCTCCGGAAGTCAGGCTTCCAGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5520
Db 5169 TCAGATTTCTCCGGAAGTCAGGCTTCCAGCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5228
QY 5521 CTCTTAGCAAACTCAGGCACCTTTACCCACATAGACCTCTGACAGAGAGCAGGCACT 5580
Db 5229 CTCTTAGCAAACTCAGGCACCTTTACCCACATAGACCTCTGACAGAGAGCAGGCACT 5288
QY 5581 TTACATGGAGTCTCTGTGGGAGAGCCATAGGCTACGGTGTAAAAAGAGCAGGGAAGTGT 5640
Db 5289 TTACATGGAGTCTCTGTGGGAGAGCCATAGGCTACGGTGTAAAAAGAGCAGGGAAGTGT 5348
QY 5641 GGCTGAGAAAGTCAGGACTTCACATAGAGCCCTAGCCACACAGCAATGACAGACAGA 5700
Db 5349 GGCTGAGAAAGTCAGGACTTCACATAGAGCCCTAGCCACACAGCAATGACAGACAGA 5408
QY 5701 TCCCTCTCTATCTCCCCCATAGAGTTTGGTTCGAC 5735
Db 5409 TCCCTCTCTATCTCCCCCATAGAGTTTGGTTCGAC 5443

RESULT 4
AAT31006
ID AAT31006 standard; DNA; 1679 BP.
XX
AC AAT31006;
XX
DT 26-SEP-1996 (first entry)
XX
XX Mouse cardiac alpha myosin heavy chain promoter.
DE
XX
KW Gene therapy; hypoxia related enhancer element; HREE; ischaemia;
XX reperfusion; promoter; alpha myosin heavy chain; alpha-MHC; ds.
XX
XX Mus sp.
XX
XX WO9620276-A1.
XX
XX 04-JUL-1996.
XX
XX 13-NOV-1995; 95WO-IB000996.
XX
```

PR 23-DEC-1994; 94US-00365486.
XX (STRI) SRI INT.
XX Webster KA, Bishopric NH, Murphy B, Laderoute KR, Green CJ;
XX WPI; 1996-321849/32.
XX Chimeric gene contg. therapeutic gene linked to HREB - partic. for
PT expressing SOD etc. in hypoxic tissue to reduce tissue injury caused by
PT ischaemia or reperfusion.
XX Disclosure; Page 50-51; 118pp; English.
XX The mouse cardiac-specific alpha-myosin heavy chain promoter (AAT31006)
CC restricts gene expression to cardiac tissue. The given sequence ends just
CC upstream of the ATG initiation codon. Chimeric constructs including this
CC promoter, plus a hypoxia response enhancer element and a therapeutic gene
CC can be used for the hypoxia-regulated, cardiac tissue-specific treatment
CC of tissue injury caused by ischaemia or reperfusion
XX
SQ Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;
Query Match 22.3%; Score 1279.8; DB 2; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;
Qy 3767 GAATTCCTTACTATCAAGGAACTGAGTGTGACCTGCACTGCAAAAGTGGATCTCCCT 3826
Db 1 GAATTCCTTACTATCAAGGAACTGAGTGTGACCTGCACTGCAAAAGTGGATCTCCCT 60
Qy 3827 AGACATCATGACTTTGTCTCTGGGAGCCAGCAGTGTGGAACTTCAGCTCTGAGAGTA 3886
Db 61 AGACATCATGACTTTGTCTCTGGGAGCCAGCAGTGTGGAACTTCAGCTCTGAGAGTA 120
Qy 3887 GGAGGCTCCCTCAGCCTGAAGTATGACAGATAGCAGGTTGAAAGGGGGAAGGAGAG 3946
Db 121 GGAGGCTCCCTCAGCCTGAAGTATGACAGATAGCAGGTTGAAAGGGGGAAGGAGAG 180
Qy 3947 CCTGGATGGAGCTGTGTGTGGAGCAGGGGACAGATATTAAGCTTGAAGAAGAGG 4006
Db 181 CCTGGATGGAGCTGTGTGTGGAGCAGGGGACAGATATTAAGCTTGAAGAAGAGG 240
Qy 4007 TGACCTTACCCAGTTTCAACTCAGCTTCAAGTATTAAGTAACTAGGTAAGGGCT 4066
Db 241 TGACCTTACCCAGTTTCAACTCAGCTTCAAGTATTAAGTAACTAGGTAAGGGCT 300
Qy 4067 GGGTAGGGAGTGTGTGAGAGCTCTCTCTCTCTGATGCTGATGCTGAGGCTTTTG 4126
Db 301 GGGTAGGGAGTGTGTGAGAGCTCTCTCTCTCTGATGCTGATGCTGAGGCTTTTG 360
Qy 4127 GGGAGGAGGAATGTGCCCAAGACTTAAAAAGGCGCATGGAGCCAGAGGGGCGAGGCAA 4186
Db 361 GGGAGGAGGAATGTGCCCAAGACTTAAAAAGGCGCATGGAGCCAGAGGGGCGAGGCAA 420
Qy 4187 CAGACCTTTTCATGGCAAACTTTGGGGCCCGTAGTATGATGATGATGATGATGATGAT 4246
Db 421 CAGACCTTTTCATGGCAAACTTTGGGGCCCGTAGTATGATGATGATGATGATGATGAT 480
Qy 4247 CGATACCTTCTTCTTCTTAAAGGAGGAGGAACTCGAGTTTACCCTCCCTATCAGTG 4306
Db 481 GGATCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 525
Qy 4307 ATAGAGAAAAGTGAAAGTCGAGTTTACCCTATCCCTATCAGTATGATAGAGAAAAGT 4366
Db 526 ----- 525
Qy 4367 CGAGTTTACCCTCCCTATCAGTATGATAGAGAAAAGTGAAAGTCGAGTTTACCCTA 4426
Db 526 ----- 525
Qy 4427 TCAGTATGATAGAGAAAAGTGAAAGTCGAGTTTACCCTATCAGTATGATAGAGAAAAGT 4486

Db 526 ----- 525
Qy 4487 GAAAGTCGAGTTTACCCTATCAGTATGATAGAGAAAAGTGAAAGTCGAGTTTACCCTA 4546
Db 526 ----- 525
Qy 4547 TCCCTATCAGTATGATAGAGAAAAGTGAAAGTCGAGTTTACCCTATCAGTATGATAG 4606
Db 526 -----GGTCCAGCAGAGGACTCCAAAT 4606
Qy 4607 TTAGGCACAGGCATATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGTGTGAGAG 4666
Db 549 TTAGGCACAGGCATATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGTGTGAGAG 608
Qy 4667 ATTTCTCCAAACCCAGGTAAAGGGAGTTTCGGGTGGGGCTCTTCAACCCACACAGACT 4726
Db 609 ATTTCTCCAAACCCAGGTAAAGGGAGTTTCGGGTGGGGCTCTTCAACCCACACAGACT 668
Qy 4727 CTCCCCACTAGAAAGGAACTGCCCTTCTCGAAGTGGGGTTCAGGCGGTGAGAGATCT 4786
Db 669 CTCCCCACTAGAAAGGAACTGCCCTTCTCGAAGTGGGGTTCAGGCGGTGAGAGATCT 728
Qy 4787 GACAGGGTGGCTTTCCACACAGCTGGGAAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAA 4846
Db 729 GACAGGGTGGCTTTCCACACAGCTGGGAAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAA 788
Qy 4847 CACTGGATGCCCTTCCCTTACGCTGTCTTCTCCATCTTCTCTGGGGATGCTCTCTCC 4906
Db 789 CACTGGATGCCCTTCCCTTACGCTGTCTTCTCCATCTTCTCTGGGGATGCTCTCTCC 848
Qy 4907 CGCTTGTGTTATCTTGGCTCTTCTGCTTTCAGCAAGATTGTCCTGCTGCTCTCTCTCA 4966
Db 849 CGCTTGTGTTATCTTGGCTCTTCTGCTTTCAGCAAGATTGTCCTGCTGCTCTCTCTCA 908
Qy 4967 TCTTCTCTACTGCTTCCGCTGCTTTCGCTTTCGCTGCTTTCGCTGCTCTCTCTCTCA 5026
Db 909 TCTTCTCTACTGCTTCCGCTGCTTTCGCTTTCGCTGCTTTCGCTGCTCTCTCTCTCA 968
Qy 5027 TTTCTCACTTCACTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTCT 5086
Db 969 TTTCTCACTTCACTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTCT 1028
Qy 5087 TCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5146
Db 1029 TCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1088
Qy 5147 TCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5206
Db 1089 TCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
Qy 5207 TAAACAATCTTCCAGTGTAGCCACAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 5266
Db 1149 TAAACAATCTTCCAGTGTAGCCACAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 1208
Qy 5267 CCCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5326
Db 1209 CCCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1268
Qy 5327 CTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5386
Db 1269 CTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1328
Qy 5387 CAGCTGTGGTCCACATCTTCAAGGATCTTCAAGGATCTTCAAGGATCTTCAAGGATCT 5446
Db 1329 CAGCTGTGGTCCACATCTTCAAGGATCTTCAAGGATCTTCAAGGATCTTCAAGGATCT 1388
Qy 5447 CTGTAGACAGCAGATCAGGATCTCCCGAGTCTTCCCGAGTCTTCCCGAGTCTTCCCG 5506
Db 1389 CTGTAGACAGCAGATCAGGATCTTCCCGAGTCTTCCCGAGTCTTCCCGAGTCTTCCCG 1448
Qy 5507 CCAGCTGCCGCGACTCTTTAGCAAACTTCCAGCAGCTTCCAGCAGCTTCCAGCAGCTT 5566
Db 1449 CCAGCTGCCGCGACTCTTTAGCAAACTTCCAGCAGCTTCCAGCAGCTTCCAGCAGCT 1508

| | | | | | | | |
|----|------|--------------|----------------|-------------------|----------------|--------|------|
| Qy | 5567 | GAGAGCAGGCAC | TTTTCATCGAGATC | TGTGGAGAGCC | ATAGGCTACGGT | TAAAGA | 5626 |
| | | | | | | | |
| Db | 1509 | GAGAGCAGGCA | CTTTACATGGAGT | CCTGTGGAGAGCC | ATAGGCTACGGT | TAAAGA | 1568 |
| | | | | | | | |
| Qy | 5627 | GGCAGGGAAGT | TGTGTAGGAAGT | CAGCACTTTCATAGAGC | CTTAGCCACACACC | AG | 5686 |
| | | | | | | | |
| Db | 1569 | GGCAGGGAAGT | TGTGTAGGAAGT | CAGCACTTTCATAGAGC | CTTAGCCACACACC | AG | 1628 |
| | | | | | | | |
| Qy | 5687 | AAATGACAGAC | AGATCCTCTATCT | CCCCCAATAGAGT | TTTGAGT | | 5731 |
| | | | | | | | |
| Db | 1629 | AAATGACAGAC | AGATCCTCTATCT | CCCCCAATAGAGT | TTTGAGT | | 1673 |

RESULT 5

RESOL 3
ABV74357

ABV74357
ID ABV74357 standard: DNA: 1679 BP.

AC ABV74357:

DT 05-FEB-2003 (first entry)

DE Mouse cardiac alpha-myosin heavy chain promoter fragment.

Therapeutic delivery system; electrical pulse generator; brain;
electrically responsive promoter; ischaemic injury; cardiac; kidney;
arterial occlusive disease; coronary arterial disease; stroke; mouse;
alpha myosin; promoter; ds.

OS Mu8 sp.

PN WO200249669-A2.

27-JUN-2002.

20-DEC-2001: 2001WO-US049845.

PR 21-DEC-2000: 2000US-0257460P.

FR 20-AUG-2001; 200108-0313926F: XX

XX
FA (MED) / MEDTRONIC INC.:

Fi
scu SA, Paqu K, Bonnet M, Bonvan MG,
XX

XX
XX
NY 100-2007-351, 100-100000

PT coupled with genetically engineered cells in mammalian tissue.

PS Disclosure; Page; 66pp + Sequence Listing; English.

The invention relates to a therapeutic delivery system comprising an electrical pulse generator operably coupled with genetically engineered cells in a mammalian tissue. The genetically engineered cells also comprise a target gene operably coupled to an electrically responsive promoter. The delivery system is useful for the production of therapeutically useful gene or protein products, in repairing tissue injury (e.g. ischaemic injury, damaged cardiac tissue, kidney tissue, brain tissue or endothelial tissue), in stimulating cells for controlled expression of therapeutically useful gene and protein sequences and for treating peripheral arterial occlusive disease, coronary arterial disease or stroke. The present sequence is that of the mouse cardiac alpha-myosin heavy chain promoter fragment, exemplary of cardiac specific promoter regions. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

SQ Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;

```
Query Match      22.3%; Score 1279.8; DB 6; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;
```

| | | | |
|----|------|--|------|
| Qy | 3767 | GAATTCCTTTACTATCAAAAGGGAAAACTGAGTGTGTGCACCTGCGAAAAGTGAGTGTCTCCCT | 3826 |
| Db | 1 | GAATTCCTTTACTATCAAAAGGGAAAACTGAGTGTGTGCACCTGCGAAAAGTGAGTGTCTCCCT | 60 |
| Qy | 3827 | AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTCTGGNACTTCAAGTCTGAGAGAGTA | 3886 |
| Db | 61 | AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACCTTCAAGTCTGAGAGAGTA | 120 |
| Qy | 3887 | GGAGGCTCCCCTCAGCCTGAAGCTATGCGATAGCCAGGGTTGAAAGGGGGAAGGAGAG | 3946 |
| Db | 121 | GGAGGCTCCCCTCAGCCTGAAGCTATGCGATAGCCAGGGTTGAAAGGGGGAAGGAGAG | 180 |
| Qy | 3947 | CCTGGAGTGGGAGCTTGTGTGAGAGCGGCGGGAAGATAATTAAAGCCTGGAAGAGAAG | 4006 |
| Db | 181 | CCTGGAGTGGGAGCTTGTGTGAGAGCGGCGGGAAGATAATTAAAGCCTGGAAGAGAAG | 240 |
| Qy | 4007 | TGACCTTTACCAGTTGTTCAACTCACCTTCAGATTAAAAATAAATGAGGTAAAGGCCCT | 4066 |
| Db | 241 | TGACCTTTACCAGTTGTTCAACTCACCTTCAGATTAAAAATAAATGAGGTAAAGGCCCT | 300 |
| Qy | 4067 | GGGTAGGGAGTGGTGTGAGACGCTCTCTCTCTCTGTGCAATGCCGTGAGGCCCTTTG | 4126 |
| Db | 301 | GGGTAGGGAGTGGTGTGAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 360 |
| Qy | 4127 | GGGAGGAGGAATCTGCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGCGGAGGCGAA | 4186 |
| Db | 361 | GGGAGGAGGAATCTGCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGCGGAGGCGAA | 420 |
| Qy | 4187 | CAGACCTTTTCATGGGCAAACTTTGGGGCCGTAGTGATCGATTGACAAGAACTCGCCAAT | 4246 |
| Db | 421 | CAGACCTTTTCATGGGCAAACTTTGGGGCCGTAGTGATCGATTGACAAGAACTCGCCAAT | 480 |
| Qy | 4247 | CGATACCTTCTCTTCTAAGCAGCAGGAGGAACTCGAGTTTTACCACTCCCTATCAGTG | 4306 |
| Db | 481 | GSATCAAGAGGAGGAGCCAGGACAGAGAGGGAAGTGGGAGGGAG----- | 525 |
| Qy | 4307 | ATAGAGAAAAGTGAAAGTCGAGTTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAAAGT | 4366 |
| Db | 526 | ----- | 525 |
| Qy | 4367 | CGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA | 4426 |
| Db | 526 | ----- | 525 |
| Qy | 4427 | TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTTACCCTCCCTATCAGTGATAGAGAAAAGT | 4486 |
| Db | 526 | ----- | 525 |
| Qy | 4487 | GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC | 4546 |
| Db | 526 | ----- | 525 |
| Qy | 4547 | TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTCGGTACCAAGAGGAGCTCCAAAT | 4606 |
| Db | 526 | -----GGTCCAGCAGAGGAGCTCCAAAT | 548 |
| Qy | 4607 | TTAGGCGAGGAGCATATGGATGGGATATAAAGGGCTGGAGCACTGAGAGCTGTGAGAG | 4666 |
| Db | 549 | TTAGGCGAGGAGCATATGGATGGGATATAAAGGGCTGGAGCACTGAGAGCTGTGAGAG | 608 |
| Qy | 4667 | ATTCTTCCAACCCAGGTAAAGGGAATTTTCGGGTGGGGCTTTTCAACCCACACAGACCT | 4726 |
| Db | 609 | ATTCTTCCAACCCAGGTAAAGGGAATTTTCGGGTGGGGCTTTTCAACCCACACAGACCT | 668 |
| Qy | 4727 | CTCCCCACCTPAGAAGGAACTGCTCTTTCTTGGAAAGTGGGGTTCCAGGCCGTGAGAGTCT | 4786 |
| Db | 669 | CTCCCCACCTPAGAAGGAACTGCTCTTTCTTGGAAAGTGGGGTTCCAGGCCGTGAGAGTCT | 728 |
| Qy | 4787 | GACAGGTGCCCTTCCACCAAGCTGGGAAAGTTCTCACTGGCAGGAGGTTTCCAAGAAA | 4846 |
| Db | 729 | GACAGGTGCCCTTCCACCAAGCTGGGAAAGTTCTCACTGGCAGGAGGTTTCCAAGAAA | 788 |

QY 4847 CACTGGATGCCCTTCCCTTACGCTGTCTTCTCCATCTTCTCCCTGGGATGCTCTCCCTCC 4906
DB |||||
QY 789 CACTGGATGCCCTTCCCTTACGCTGTCTTCTCCATCTTCTCCCTGGGATGCTCTCCCTCC 848
DB |||||
QY 4907 CGCTGTGGTTTATCTTGTGCTCTTGTCTTTCAGCAAGATTGCTCTGTCCTCACTCCA 4966
DB |||||
QY 849 CGCTGTGGTTTATCTTGTGCTCTTGTCTTTCAGCAAGATTGCTCTGTCCTCACTCCA 908
DB |||||
QY 4967 TCTTCTCTACTGCTCGTGGCTTGTGCTTCTGCTTCTGCTGCTTCTTCTTCCACCCA 5026
DB |||||
QY 909 TCTTCTCTACTGCTCGTGGCTTGTGCTTCTGCTTCTGCTGCTTCTTCTTCCACCCA 968
DB |||||
QY 5027 TTTCTCACTTCACTTCTTCTCCCTTCTCACTTGTGATTCATCTTCTTCTTCTTCTTCT 5086
DB |||||
QY 969 TTTCTCACTTCACTTCTTCTCCCTTCTCACTTGTGATTCATCTTCTTCTTCTTCTTCT 1028
DB |||||
QY 5087 TCTTCT 5146
DB |||||
QY 1029 TCTTCT 1088
DB |||||
QY 5147 TCTTCT 5206
DB |||||
QY 1089 TCTTCT 1148
DB |||||
QY 5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTCTGCTGGTGTCTCTTACCTTCTCTCA 5266
DB |||||
QY 1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTCTGCTGGTGTCTCTTACCTTCTCTCA 1208
DB |||||
QY 5267 CCCCTGCTGTCT 5326
DB |||||
QY 1209 CCCCTGCTGTCT 1268
DB |||||
QY 5327 CTACT 5386
DB |||||
QY 1269 CTACT 1328
DB |||||
QY 5387 CAGCTGTCT 5446
DB |||||
QY 1329 CAGCTGTCT 1388
DB |||||
QY 5447 CTGTAGACAGAGATCAGATTTCTCCCGAAGTTCAGGCTTCCAGCCCTCTCTCTCTCTGC 5506
DB |||||
QY 1389 CTGTAGACAGAGATCAGATTTCTCCCGAAGTTCAGGCTTCCAGCCCTCTCTCTCTCTGC 1448
DB |||||
QY 5507 CCAGTGTCCCGGACTTTAGAAACCTTCAGGACCTTACCCACATAGACCTCTGACA 5566
DB |||||
QY 1449 CCAGTGTCCCGGACTTTAGAAACCTTCAGGACCTTACCCACATAGACCTCTGACA 1508
DB |||||
QY 5567 GAGAAGCAGGCACTTTACATGAGTCTCTGGTGGGAGGCCATAGGCTACGGTGTAAAGA 5626
DB |||||
QY 1509 GAGAAGCAGGCACTTTACATGAGTCTCTGGTGGGAGGCCATAGGCTACGGTGTAAAGA 1568
DB |||||
QY 5627 GGCAGGGAAGTGTGGTGTAGGAAAGTTCAGGACTTTCACATAGAGCCCTAGCCACACAG 5686
DB |||||
QY 1569 GGCAGGGAAGTGTGGTGTAGGAAAGTTCAGGACTTTCACATAGAGCCCTAGCCACACAG 1628
DB |||||
QY 5687 AAATGACAGACAGATCCCT 5731
DB |||||
QY 1629 AAATGACAGACAGATCCCT 1673
DB |||||

RESULT 6

ADL14241
ID ADL14241 standard; DNA; 1679 BP.
XX
AC ADL14241;
XX
XT AC
DT 06-MAY-2004 (first entry)
XX
DE Mouse cardiac alpha-myosin heavy chain promoter region.
XX
KW Mouse; cardiac alpha-myosin heavy chain promoter; ds;
KW electrical pulse generator; electrically responsive promoter;

electrical response enhancer element; pacemaker.

Mus sp.

US2003204206-A1.

30-OCT-2003.

20-DEC-2001; 2001US-00027655.

21-DEC-2000; 2000US-0257460P.

20-AUG-2001; 2001US-0313926P.

(MEDT) MEDTRONIC INC.

Padua RA, Schu CA, Bonner MD, Donovan MG, Soykan O;

WPI; 2004-032680/03.

Therapeutic delivery system useful for regulating delivery of therapeutic proteins and nucleic acids, comprises electrical pulse generator coupled with genetically engineered cells in mammalian tissue.

Disclosure; SEQ ID NO 5; 39pp; English.

The invention relates to a therapeutic delivery system comprising an electrical pulse generator coupled with genetically engineered cells in mammalian tissue. The genetically engineered cells further include a target gene coupled to an electrically responsive promoter. The invention also relates to an expression vector comprising an electrical response enhancer element, a tissue specific promoter heterologous to the element and a coding sequence, an apparatus for testing cells comprising an upper plate electrode, a lower plate electrode and a porous membrane positioned between electrodes during operation, and a method of treating a patient comprising providing the patient with an electrical pulse generator coupled with genetically engineered cells in a patient tissue. The electrical pulse generator is a pacemaker. The method is used for regulating the delivery of therapeutic proteins and nucleic acids. The invention provides controlled and local delivery of therapeutically important gene or protein products. This sequence represents the mouse cardiac alpha-myosin heavy chain promoter region, used in the method of the invention.

Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;

Query Match 22.3%; Score 1279.8; DB 12; Length 1679;

Best Local Similarity 82.2%; Pred. No. 0;

Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

QY 3767 GAATTCCTTACTATCAAGGAACTGAGTCGTGCACCTGCAAGTGATGCTCTCCCT 3826

DB 1 GAATTCCTTACTATCAAGGAACTGAGTCGTGCACCTGCAAGTGATGCTCTCCCT 60

QY 3827 AGACATCATGACTTTGTCTCTGGGAGCCAGCACCTGTGAACTTCAGGCTCTGAGAGATG 3886

DB 61 AGACATCATGACTTTGTCTCTGGGAGCCAGCACCTGTGAACTTCAGGCTCTGAGAGATG 120

QY 3887 GGAGGCTCCCTCAGCCCTGAAGCTATGAGATAGCCAGGTTGAAAGGGGAGGAGAG 3946

DB 121 GGAGGCTCCCTCAGCCCTGAAGCTATGAGATAGCCAGGTTGAAAGGGGAGGAGAG 180

QY 3947 CCTGGATGGAGCTTGTGTGGAGCAGGGGACAGATATTAAGCTTGAAGAGAGAG 4006

DB 181 CCTGGATGGAGCTTGTGTGGAGCAGGGGACAGATATTAAGCTTGAAGAGAGAG 240

QY 4007 TGACCCCTTACCAGTGTTCACCTCACCTTCAGATTAATAAATACTCAGGTAGGGCT 4066

DB 241 TGACCCCTTACCAGTGTTCACCTCACCTTCAGATTAATAAATACTCAGGTAGGGCT 300

QY 4067 GGGTAGGGAGGTGTGTGAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4126

DB 301 GGGTAGGGAGGTGTGTGAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360

CC promoter, which is an atria-preferred promoter. The promoter can be used
 CC in compositions of the invention for altering cardiac-preferred
 CC expression in transgenic animals. A claimed transgenic rabbit has in its
 CC genome a stably incorporated expression cassette comprising: a promoter
 CC having e.g. the present sequence or at least 90% identity to it or
 CC comprising at least 50 contiguous nucleotides of it capable of initiating
 CC transcription in an animal cell; and a heterologous nucleotide sequence
 CC that is operably linked to the promoter. In the present case, the
 CC promoter is capable of initiating atria-preferred transcription. Other
 CC transgenic animals, e.g. mouse, dog, pig, goat, cow, monkey, chimpanzee
 CC and sheep, may also be produced. The transgenic animal exhibits altered
 CC cardiac preferred expression of the heterologous nucleotide sequence and
 CC may exhibit an altered susceptibility to cardiomyopathy (or cardiomyopathy).
 CC Such animals can be used to identify anti-cardiopathic compounds.
 XX

Sequence 5190 BP; 1095 A; 1525 C; 1642 G; 928 T; 0 U; 0 Other;

Query Match 7.1%; Score 407.6; DB 13; Length 5190;
 Best Local Similarity 58.8%; Pred. No. 1.5e-104;
 Matches 1265; Conservative 0; Mismatches 699; Indels 188; Gaps 25;

| | | | |
|----|------|--|------|
| Qy | 2099 | ACCTGTGGCCAGAGGCTCAT-CTGTGTCCACATCTCTAGAAATGTTTCATCAGACTGCAG | 2157 |
| Db | 2140 | ACCTTTGGCCAGAGTATCCGTCGCGCTCCACACTCCAGGGCTCCAGCAGAAATGCGG | 2199 |
| Qy | 2158 | GGCTGGCTTGGAGGSCAGCTG-----GAAAGATATGTGAGAGCCAGGG | 2201 |
| Db | 2200 | AGCTGATTTGGGTGGCAGTGAGACAGCGTGGACAGAGAGGTATGGGAACGCCAGGA | 2259 |
| Qy | 2202 | GAGACAAGGGGGCTTAGGAAAGGAAGAGGGGCAAAACAGGCCACACAGAGGGCAGAG | 2261 |
| Db | 2260 | CAGAGGAGGGGGCCAGGGAAG--GAAGAGGGCAAAACAGGCCAGACCG--GGCCCGAG | 2314 |
| Qy | 2262 | CCGAGAACTAGATTAACTCTCTCTTGTGTGATCTTCCATAGAGGCGAGTGGGAACTCTG | 2321 |
| Db | 2315 | CCGAGAGCAGGCTAAACCCCTTCTTCTTACCTCTTCCCTAGGAGG--GCAGGCACTGCC | 2372 |
| Qy | 2322 | TGACCACCATCCCATGAGCCCCACCTACCCATACCAAGTTTGGCCCTGAGTGGCATTTCT | 2381 |
| Db | 2373 | TGCCACCCCTCAGCCCGCCCTTACCACCTCTCCCAAGTTTGGTCTGCTCAGGGGCGCA | 2432 |
| Qy | 2382 | AGGTTTC-CCTGAGGACAGAGCTGGCCCTTGTCTCTTGGACCTGACCCAAAGCTGACCCAA | 2440 |
| Db | 2433 | GGGTTCTCCCGAGGACAAAGCTGGTCTTTGTACCTGGACCTTACCCAGGCTGACCTGG | 2492 |
| Qy | 2441 | TGTTCTCAGTACCTTTGAATGCCCTCAAGAGCTTGAGAACAGGACAGGAGTGACATATTAGGC | 2500 |
| Db | 2493 | TATTC-CAGGCCCTTATCAGGTCCCGAGGGGATGGGAACAGGCGAGTGACGCCACCAAGGC | 2551 |
| Qy | 2501 | CATGGCTAACCTTGGAGCTTGACACAGGAGCCCTCAAGTGACCTCCAGGGACACAGCTG | 2560 |
| Db | 2552 | CACGGGATAACCC-CCAGCTTGTGCACCTGGGGCCCTCAAGTGACCT-----AGAGCTA | 2602 |
| Qy | 2561 | CAGACAGTGGCCCTTATCCCAAGAGACCAACATTTGGCATAGTGCTGCTGCTG--AAATG | 2617 |
| Db | 2603 | GAGACAGTGGCCCTGTCATCCCTGAGGGCAACGGTTTGGCATGTGCACCTGCAGGAGATG | 2662 |
| Qy | 2618 | GGAAATGCAAGTGTGAATCAGGTCCCTTCAAGAAATACATGCATGCAAGACCTTAAGACCCCTG | 2677 |
| Db | 2663 | GGAAATGCGCGGTCAACTCAGGTCCCTTCAAGGACACTGAGTCTGGGGTCCCAAGTCCCCCA | 2722 |
| Qy | 2678 | GAGAGAGGGGTATGCTCTCTCCCCCAACCCACCAATGAAGGGGAGTGAATCTCTAGGGGGC | 2737 |
| Db | 2723 | GAGAGAGATGTGTCTCTGAGCCCG------TAATAGGGAGCGCAGCATCTAGGGGGC | 2774 |
| Qy | 2738 | TGGGACCTTGGGAGACACCA--CATTTACTGAGAGTCTGAG-CCCAAGAAAACCTGACC | 2794 |
| Db | 2775 | TGACAACCTTGGGGGAACCCCACTTCCAGGCAAGTGTCTCCACCCAGAGAGAGCTGACC | 2834 |
| Qy | 2795 | GCCCTGTGTCTGCCACCTCCACACTCTAGAGCTATATTGAGAGGTCACAGTAGATAGG | 2854 |
| Db | 2835 | CCTGCCCTCTACCCACCTCCACACCTCTGGAGCTATATCGAGAGGTGTCTAGTGGATAGG | 2894 |

| | | | |
|----|------|---|------|
| Qy | 2855 | GTGGAGCTGCTAGCAGGAGAGTGTCTTCTGGTGTGAGGGTGTAGGGGAAGCCAGAGC | 2914 |
| Db | 2895 | G-GTGGAGGGAGCTGAGGCAGTATTTCTTGGTGTGAGGGTGTGGGGAAAGCCAGGC | 2953 |
| Qy | 2915 | AGGGAGTCTGGCTTTGTCTCTGAACAACAATGTCTACTTATTATAACAGGCATGACCT | 2974 |
| Db | 2954 | AGGAGAGTCTGGCTTTGTCTTCTGAACAACAATGTCCAATAGTCAACAAGGCACGGCT | 3013 |
| Qy | 2975 | GCTAAGACCCCAACATCTACGACCTCTGAAA--AGACAGCAGCCCTGAGGACAGGGTT | 3032 |
| Db | 3014 | GTTGAAGACCCGACACCTTACCGCTCTGAGAGGGGACAGCCTGCGAGGATGGAGGAGT | 3073 |
| Qy | 3033 | GTCTCTGAGCCTTGGGTCTTGTATGGTCCCAAAAGAGGGCATGAGTGTGAGTATAAGG | 3092 |
| Db | 3074 | ACCTTGAGGCTGGGTGCTTGAGGACCCCCCCCCCAAGACAGCATGGGAGTGTGAGC | 3133 |
| Qy | 3093 | CCCAGAGAGCTTAGAAGAGGGCACTTGGGAAGGGTCACTGCTGAGAGCCCCCTATTCAT | 3152 |
| Db | 3134 | TCAGAGGGGATAGCAGGCTTCTATCTGGGAAGGGGCGAGACTACAGAGCCCCCTCCCC | 3193 |
| Qy | 3153 | GGAACTCTGAGCCTGGGGCCAACTGGTGTAAATCTCTGGGCTTCCAGGCATTCAAGCA | 3212 |
| Db | 3194 | CAACCTCTGCACCCCCCTGSCAGCCCTGTGGAAACGCGAGAGCCCAATTGGTGTAAAT | 3253 |
| Qy | 3213 | GCACCTGCATCTCTGGCAGCTTGGGAGGGAGGAGCAACCCCCACCTTATACCT | 3272 |
| Db | 3254 | CCCAGGCTGTCTGGGCCCCAGGCATCTCTGGCAGCCTCTGACCTCTGGCAGCTG | 3313 |
| Qy | 3273 | TTCTCCCTCAGCCCCAGGATTAACACCTCTGGCCTTCCCCCTTCCACCTCCCATCAGGA | 3332 |
| Db | 3314 | GAGAGGGGAAGGAGCGCCCCCACCCCTCTGCTCTCCTCAGGCCCGGGATTAGT | 3373 |
| Qy | 3333 | GTGAGGGTTTCAGAGGGAGGGTAAACCTTACATGTCTCAACATCATGTGTGCGACATAT | 3392 |
| Db | 3374 | GCAGAAGGTGCAGCAGGAGGGTAAAGCCCTTTGGTGCA-----GATGT | 3418 |
| Qy | 3393 | ATGATCAGTATGTGTAGAGGCAAGAAAGAAATCTGCAGGCTTAACCTGGGTAAATGTGT | 3452 |
| Db | 3419 | GAGCATCGGCTGTGGAAAGACGCGAGGATCTCTGTAACTGAAGTGTGAAA----- | 3472 |
| Qy | 3453 | AAAGTCTGTGTGCATGT | 3512 |
| Db | 3473 | -----TCGTCATGTGTCTGGTCA | 3492 |
| Qy | 3513 | GTTCTGTGTGTGAGTTACACAGCTGCAGGTTTGTGTGTAAATTTGCCCAAGCAAGTGG | 3572 |
| Db | 3493 | AGTCTGCACGCGAGGGTCTGCAAGGGTCTGTGCGCAGGTGTGCTGAGGCGACGCTG | 3552 |
| Qy | 3573 | GTGAATCCCTTCCATGTTTAAAGAGATTGATGAGTGGCTTCTCAAGSACCATGGA | 3632 |
| Db | 3553 | GAGAAATTCCTTCCACTGGGTGATGGTTTGCATGAGGGCCACTGTATGGGAGACCCGG | 3612 |
| Qy | 3633 | AAATAGAATGACACTCTATATGTGTCTTAAGCTTAAGGTAGCAAGGCTTTTGGAGGACA | 3692 |
| Db | 3613 | GAGTGTGGGGCTCCAGATGCATTCCCAAGCAAGAGACCTTGAGTCTCTGTGACACA | 3672 |
| Qy | 3693 | CCTGTCTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGAGTAAGAGAGA | 3752 |
| Db | 3673 | CCTGCC-----TGGGAATAGAAGCCAGGCCACAGCGTGTGTGCAGAGTTGGGGCAGG | 3724 |
| Qy | 3753 | GAGAGGGGGTGTAGAATTTCTTACTATCAAGGGGAACCTGAGTCTGACCTGCAAG | 3812 |
| Db | 3725 | ACAGAG-----AGGGAACTGAGTC-TGGAGCTGGCTAG | 3757 |
| Qy | 3813 | TGGATGTCTCTCCCTAGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCA | 3872 |
| Db | 3758 | TAGAGCTGTCTCAGCTCACAGG-----TCCA | 3785 |
| Qy | 3873 | GGTCTGAGAGATGAGGAGCTCCCTCAGCTT-CAGAGTATGCAGATAGCCAGGGTTGAA | 3931 |
| Db | 3786 | GGTCTGAGAGTGTGGGAGACACTTTCAGCTCTGGCCTGTGTGGAAGACAGGCTTGGGTA | 3845 |
| Qy | 3932 | AGGGGAAGGA-GAGCCTGGGATGGGAGCTTTGTGTGTGGAGGACAGGGAACAGATATTA | 3990 |

```
Db 3846 TGGGAGAGGGTAGGGCAAGGGAGGAACTGCAGCTGGGGGGCAGGGGCAAGCAATTC 3905
Qy 3991 AGCCTGGAGAGAGGTGACCTTACCCAGTGT--TCAACTCACCTTCAGATTAAAA 4048
Db 3906 GTCCCTATATGAAAGGTGACCTTACCCAGTGTCTCAACTCACCTTCAGGTTAAAA 3965
Qy 4049 TAACTGAGGTAAAGGCTCTGGGTAGGGGAGGTGGTCT---GAGAGCTCTCTGTCTCTCT 4104
Db 3966 TAAACGAGGTAAAGGCTATGTGGGCGGGGAGGTGTGGAGAGGTCTCTGTCTTCCCA 4025
Qy 4105 CTGCATGCCCTGAGGCCCTTTTGGGAGGAGGAGTAATGTGCCCAAGGACTAAAAAAGGCCAT 4164
Db 4026 CTATCTGCCCATCAGCGCTCTGGAGGGGCGGAATGTCTCAAGGACTAAAAAAGGCCCT 4085
Qy 4165 GGAGCCAGGGGCGAGGGCAACAGACTTTCATGGGCAACCTTGGGGGCC 4216
Db 4086 GGAGCCGAGGGCTGGGGCGAGCAGACCTTTTCATGGGCAAACTCTGGGGGCC 4137

RESULT 8
AAT11359
ID AAT11359 standard; DNA; 520 BP.
AC AAT11359;
XX
XX 07-JUL-1996 (first entry)
XX
XX Bidirectional promoter.
XX
XX Tet repressor; Herpes simplex virus; HSV; virion protein 16; VP16;
XX fusion protein; gene expression; regulation; inhibition; activation;
XX transcription; ds.
XX
XX Synthetic.
XX
XX WO9601313-A1.
XX
XX 18-JAN-1996.
XX
XX 29-JUN-1995; 95WO-US008179.
XX
XX 01-JUL-1994; 94US-00270637.
XX 15-JUL-1994; 94US-00275876.
XX 03-FEB-1995; 95US-00383754.
XX 07-JUN-1995; 95US-00486814.
XX
XX (BUJA/) BUJARD H.
XX (GOSS/) GOSSSEN M.
XX
XX Bujard H, Gossen M;
XX
XX WPI; 1996-087666/09.
XX
XX New tetracycline-regulated transcription modulators - comprising fusion
XX proteins which bind to tet operator sequences to activate or inhibit
XX transcription.
XX
XX Disclosure; Page 73; 112pp; English.
XX
XX Fusion proteins comprising a first polypeptide which binds to a tet
XX operator sequence in the presence of tetracycline or a tetracycline
XX analogue, operatively linked to a second polypeptide which either
XX activates or inhibits transcription in eukaryotic cells may be used to
XX activate or inhibit transcription. Such proteins may be used to regulate
XX gene expression in cells and may be particularly useful for gene therapy
XX and for expression of gene products in transgenic organisms. Induction of
XX gene expression is rapid, efficient and strong, typically 1000-2000 fold.
XX The inducing agent does not cause pleiotropic effects or cytotoxicity in
XX eukaryotic cells. This sequence is a bidirectional promoter which can be
XX used in the production of vector constructs. See also AAT11358
XX
XX Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
```

```
Query Match 5.4%; Score 311.8; DB 2; Length 520;
Best Local Similarity 99.4%; Pred. No. 8.7e-78;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4277 GGAACCTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCAC 4336
Db 63 GGATCTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCAC 122
Qy 4337 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182
Qy 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 242
Qy 4457 TACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 4516
Db 243 TACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 302
Qy 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGT 4576
Db 303 ATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGT 362
Qy 4577 CGAGCTCGGTACCAG 4591
Db 363 CGAGCTCGGTACCCG 377

RESULT 9
AAT45722
ID AAT45722 standard; DNA; 520 BP.
XX
XX AC AAT45722;
XX
XX 18-MAR-1997 (first entry)
XX
XX Bi-directional tetracycline-regulated promoter region.
XX
XX Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
XX repressor; gene expression; therapy; transgenic animal; disease model;
XX HSV; herpes simplex virus; tk; thymidine kinase; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX promoter 24..427
XX /*tag= a
XX /note= "bi-directional, i.e. translation starts in a 3'
XX to 5' direction at posn. 24 and translation starts in 5'
XX to 3' direction at posn. 427, an explanatory figure is
XX given in the specification"
XX repeat_region 69..363
XX /*tag= b
XX /rpt_type= direct
XX /note= "seven repeats of the tet operator sequence
XX contained within the bi-directional promoter"
XX
XX WO9640892-A1.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-US009049.
XX
XX 07-JUN-1995; 95US-00485971.
XX (BADI ) BASF AG.
XX
XX Bujard H, Gossen M, Hillen W, Helbl V, Schnappinger D;
XX WPI; 1997-052305/05.
XX
```


expression in a cell, using tetracycline-regulated fusion proteins. The method involves obtaining a cell from a subject, and introducing a nucleic acid molecule into the cell, which operatively links a gene to at least one tetracycline (tet) operator sequence. A second nucleic acid molecule is then introduced which encodes a fusion protein, where the fusion protein comprises a first polypeptide which binds to a tet operator sequence in the presence of tetracycline, or its analogue, operatively linked to a second polypeptide (e.g. VP16) which activates transcription in eukaryotic cells to form a modified cell. The modified cell can then be administered to the subject, and the concentration of tetracycline (or and analogue) can be regulated so that the expression of the gene is regulated. The method is useful for gene therapy of diseases such as cancer and arthritis or for tissue regeneration and wound healing. The method may also be useful for the production of proteins in vitro and in vivo and for the production of stable cell lines for cloning

XX Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match 5.4%; Score 311.8; DB 3; Length 520;
 Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
 DB 63 GGATCTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 122

QY 4337 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
 DB 123 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182

QY 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4456
 DB 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 242

QY 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 4516
 DB 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 302

QY 4517 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4576
 DB 303 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 362

QY 4577 CGAGTCGGTACCAG 4591
 DB 363 CGAGTCGGTACCAG 377

RESULT 13

ID AAH47631

XX AAH47631 standard; DNA; 520 BP.

AC AAH47631;

XX 30-NOV-2001 (first entry)

DT 30-NOV-2001 (first entry)

XX Nucleotide sequence of a bidirectional promoter region.

DE Fusion protein; tet operator; tetracycline; gene therapy; anti-cancer;

KW rheumatoid arthritis; hypopituitarism; wound healing; hemophilia;

KW diabetes; Alzheimer's disease; tet repressor; promoter; ds.

XX Unidentified.

OS US6271348-B1.

XX 07-AUG-2001.

XX 24-JAN-2000; 2000US-00489777.

XX 14-JUN-1993; 93US-00076726.

XX 19-JUN-1993; 93US-00076327.

XX 14-JUN-1994; 94US-00260452.

XX 01-JUL-1994; 94US-00270637.

PR 15-JUL-1994; 94US-00275876.
 PR 03-FEB-1995; 95US-00383754.
 PR 07-JUN-1995; 95US-00485978.
 PR 28-SEP-1998; 98US-00162184.

XX (BADI) BASF AG.
 PA (KNOL) KNOLL AG.

XX Bujard H, Goessen M;

PI WPI; 2001-556625/62.

XX Fusion protein for inhibiting transcription in eukaryotic cells useful in

XX gene therapy applications comprises a first polypeptide, which binds to

XX tet operator sequences, operatively linked to a heterologous second

XX polypeptide.

PS Disclosure; Fig 7B; 69pp; English.

XX The invention relates to a fusion protein that comprises a first

XX polypeptide which binds to tet operator sequences, operatively linked to

XX a heterologous second polypeptide, which inhibits transcription in

XX eukaryotic cells. The fusion proteins are tetracycline-responsive and are

XX useful for regulation of transcription in eukaryotic cells and animals.

XX The tetracycline (Tc)-controlled regulatory system is useful in various

XX applications in gene therapy, such as in the treatment of various disease

XX conditions e.g. rheumatoid arthritis, hypopituitarism, wound healing and

XX tissue regeneration, anticancer treatments, benign prostatic hyper trophy,

XX hemophilia, diabetes and arteriosclerosis. They are also useful for bone

XX marrow support therapy, treatment of central nervous system disorders

XX e.g. Alzheimer's disease, Parkinson's disease (see AAH47628 for a

XX detailed description of the uses). The present sequence represents the

XX nucleotide sequence of a bidirectional promoter region for coordinate

XX regulation of two genes of interest by tetracycline- regulated

XX transcriptional activator

XX SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match 5.4%; Score 311.8; DB 4; Length 520;
 Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 4336

DB 63 GGATCTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 122

QY 4337 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396

DB 123 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182

QY 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4456

DB 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 242

QY 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 4516

DB 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 302

QY 4517 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4576

DB 303 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 362

QY 4577 CGAGTCGGTACCAG 4591

DB 363 CGAGTCGGTACCAG 377

RESULT 14

AAH25572

ID AAH25572 standard; DNA; 520 BP.

XX AC AAH25572;

XX

DT 05-SEP-2001 (first entry)
 XX Nucleotide sequence of a bidirectional promoter.
 DE Transgenic plant; transgene; tet operator-linked gene; Tet repressor;
 KW tetracycline-regulated transcriptional regulatory system; tet operator;
 KW tetracycline; ds.
 XX Unidentified.
 XX US6242667-B1.
 PN 05-JUN-2001.
 XX 28-SEP-1998; 98US-00161902.
 XX 14-JUN-1993; 93US-00076327.
 PR 14-JUN-1993; 93US-00076726.
 PR 14-JUN-1994; 94US-00260452.
 PR 01-JUL-1994; 94US-00270637.
 PR 15-JUL-1994; 94US-00275876.
 PR 03-FEB-1995; 95US-00383754.
 PR 07-JUN-1995; 95US-00487472.
 XX (BADI) BASF AG.
 PA (KNOL) KNOLL AG.
 XX Bujard H, Gossen M;
 XX WPI; 2001-396837/42.
 DR Transgenic plant for regulating the expression of genes, comprises a
 PT transgene integrated into the genome and a tet operator-linked gene in
 PT the genome.
 XX Disclosure; Fig 7B; 65pp; English.
 XX The specification describes a transgenic plant with a tetracycline-
 CC regulated transcriptional regulatory system. The transgenic plants have a
 CC transgene and a tet operator-linked gene integrated in the genome, which
 CC confers a detectable and functional phenotype on the plant when
 CC expressed. The transgene comprises a transcriptional regulatory element
 CC functional in cells of the plant operatively linked to a polynucleotide
 CC sequence encoding a fusion protein that activates transcription of the
 CC tet operator linked gene. The fusion protein comprises a first
 CC polypeptide that is a mutated Tet repressor that binds to a tet operator
 CC sequence in the presence of tetracycline or its analogue, operatively
 CC linked to a second polypeptide that activates transcription in eukaryotic
 CC cells. In the presence of tetracycline, the fusion protein binds to the
 CC tet operator-linked gene and activates transcription of the tet operator-
 CC linked gene such that it is expressed at a level sufficient to confer the
 CC detectable and functional phenotype on the plant. The level of expression
 CC of the tet-operator gene can be downmodulated by depleting tetracycline
 CC from the plant. The transgene can be used to regulate the expression of
 CC genes in the plant. The transgenic plant can be used to analyse the
 CC functions of cellular proteins. The present sequence represents a
 CC bidirectional promoter for coordinate regulation of two genes of interest
 CC by a tetracycline-regulated transcriptional activator. It is used to
 CC produce transgenic plants of the invention
 XX
 SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
 Query Match 5.4%; Score 311.8; DB 5; Length 520;
 Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4277 GGAATCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
 DB 63 CGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 122
 QY 4337 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
 DB 123 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

QY 4397 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4456
 DB 183 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 242
 QY 4457 TACCACCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
 DB 243 TACCACCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
 QY 4517 ATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGT 4576
 DB 303 ATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGT 362
 QY 4577 CGAGCTCGGTACCAG 4591
 DB 363 CGAGCTCGGTACCCG 377
 RESULT 15
 ACA94734
 ID ACA94734 standard; DNA; 520 BP.
 XX ACA94734;
 AC ACA94734;
 XX 11-AUG-2003 (first entry)
 DT Tet coordinated bi-directional promoter region #2.
 DE Gene therapy; tet promoter; transgenic; rheumatoid arthritis; ds;
 KW hypopituitarism; wound healing; anti-cancer treatment; promoter;
 KW transgenic farm animal; stable cell line production; tetracycline.
 XX Unidentified.
 OS US2003022315-A1.
 PN 30-JAN-2003.
 PD 03-AUG-2001; 2001US-00921650.
 PF 14-JUN-1993; 93US-00076327.
 PR 01-JUL-1994; 94US-00270637.
 PR 15-JUL-1994; 94US-00275876.
 PR 03-FEB-1995; 95US-00383754.
 PR 07-JUN-1995; 95US-00485978.
 PR 28-SEP-1998; 98US-00162184.
 PR 24-JAN-2000; 2000US-00489777.
 XX (BADI) BASF AG.
 PA (KNOL) KNOLL AG.
 XX Bujard H, Gossen M;
 XX WPI; 2003-438975/41.
 DR Regulating expression of a gene in cell of a subject, by utilizing
 PT components of tetracycline repressor/operator inducer system of
 PT prokaryotes to regulate gene expression in eukaryotic cells.
 XX Disclosure; Fig 7B; 71pp; English.
 XX The invention relates to a method of regulating expression of a tet
 CC operator-linked gene in cell of a subject, which involves introducing
 CC into the cell a nucleic acid molecule encoding a fusion protein which
 CC inhibits transcription in eukaryotic cells, comprising a first
 CC polypeptide which binds to a tet operator sequence, operatively linked to
 CC a heterologous polypeptide which inhibits transcription in eukaryotic
 CC cells and modulating concentration of tetracycline, or its analogue in
 CC the subject. The method is useful for regulating expression of a
 CC exogenous or endogenous gene in a cell. The system has widespread
 CC applicability to the study of cellular development and differentiation in
 CC eukaryotic cells, plants and animals. For expression of e.g. oncogenes
 CC can be regulated in a controlled manner in cells to study their function.

CC The system can be used to regulate the expression of site-specific
CC recombinases such as CRE or FLP, to allow for irreversible modification
CC of the genotype of a transgenic organism under controlled conditions at a
CC particular stage of development. The method is also useful for gene
CC therapy purposes, in treatment for either genetic or acquired diseases,
CC including rheumatoid arthritis, hypopituitarism, wound healing and anti-
CC cancer treatments, for large-scale production of proteins in vitro and in
CC transgenic farm animals and for production of stable cell lines for gene
CC cloning. The present sequence represents DNA encoding a tet promoter.
XX
SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match 5.4%; Score 311.8; DB 8; Length 520;
Best Local Similarity 99.4%; Pred. No. 8.7e-78;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4277 GGAATCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 122
QY 4337 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
DB 123 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182
QY 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4456
DB 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 242
QY 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 4516
DB 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 302
QY 4517 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGT 4576
DB 303 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCAG 4591
DB 363 CGAGCTCGGTACCG 377

Search completed: April 25, 2005, 03:43:34
Job time : 1993.22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 25, 2005, 02:20:37 ; Search time 12059.9 Seconds
(without alignments)
16101.184 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatccgcaaggtcacaca.....ccataagagtttgatcgac 5735

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsal:*

9: gb_gsal2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| C 1 | 534.4 | 9.3 | 613 | 8 | AZ269661 RPCI-23-4 |
| C 2 | 499.4 | 8.7 | 529 | 8 | BH073204 RPCI-24-3 |
| C 3 | 473.8 | 8.3 | 496 | 8 | AZ300332 RPCI-23-1 |
| C 4 | 305.6 | 5.3 | 312 | 8 | AZ457684 IM0261812 |
| C 5 | 305.2 | 5.3 | 469 | 8 | BH285252 CH230-118 |
| C 6 | 266.4 | 4.6 | 447 | 8 | AZ729312 RPCI-24-7 |
| C 7 | 236 | 4.1 | 447 | 5 | BY000251 BY000251 |
| C 8 | 225.2 | 3.9 | 616 | 2 | BF525096 UI-R-ACO |
| C 9 | 145.4 | 2.5 | 370 | 8 | AZ407663 IM0178C01 |
| C 10 | 133.4 | 2.3 | 390 | 5 | BY000731 BY000731 |
| C 11 | 133.4 | 2.3 | 409 | 5 | BY061731 BY061731 |
| C 12 | 133.4 | 2.3 | 415 | 5 | BY083328 BY083328 |
| C 13 | 133.4 | 2.3 | 739 | 6 | CB596657 AGENCOURT |
| C 14 | 133.4 | 2.3 | 741 | 6 | CB599099 AGENCOURT |
| C 15 | 133.4 | 2.3 | 783 | 6 | CA490238 AGENCOURT |
| C 16 | 133.4 | 2.3 | 841 | 6 | CA460851 AGENCOURT |
| C 17 | 133.4 | 2.3 | 843 | 6 | CA495692 AGENCOURT |
| C 18 | 133.4 | 2.3 | 850 | 6 | CA490440 AGENCOURT |
| C 19 | 133.4 | 2.3 | 917 | 6 | CA490678 AGENCOURT |
| C 20 | 133.4 | 2.3 | 920 | 6 | CA490677 AGENCOURT |
| C 21 | 132.8 | 2.3 | 871 | 6 | CA490555 AGENCOURT |
| C 22 | 132.4 | 2.3 | 363 | 5 | BY083663 BY083663 |
| C 23 | 131.8 | 2.3 | 881 | 6 | CB951696 AGENCOURT |
| C 24 | 130 | 2.3 | 253 | 5 | BX632021 BX632021 |

| | | | | | |
|------|-------|-----|-----|---|----------|
| C 25 | 130 | 2.3 | 459 | 1 | AI642327 |
| C 26 | 129.6 | 2.3 | 378 | 5 | BY064144 |
| C 27 | 129.2 | 2.3 | 224 | 5 | BY000659 |
| C 28 | 129 | 2.2 | 707 | 1 | AA855323 |
| C 29 | 127 | 2.2 | 349 | 1 | AI892571 |
| C 30 | 126.8 | 2.2 | 412 | 5 | BY435020 |
| C 31 | 126.8 | 2.2 | 468 | 5 | BY414811 |
| C 32 | 126.2 | 2.2 | 243 | 5 | BY000397 |
| C 33 | 125.2 | 2.2 | 368 | 2 | BF472804 |
| C 34 | 125.2 | 2.2 | 385 | 5 | BY428669 |
| C 35 | 125.2 | 2.2 | 428 | 2 | BF463620 |
| C 36 | 124.8 | 2.2 | 399 | 5 | BY432934 |
| C 37 | 123.8 | 2.2 | 405 | 5 | BY419314 |
| C 38 | 123.8 | 2.2 | 425 | 5 | BY427154 |
| C 39 | 123.6 | 2.2 | 405 | 5 | BY436012 |
| C 40 | 123 | 2.1 | 426 | 1 | AA140343 |
| C 41 | 123 | 2.1 | 473 | 1 | AA919406 |
| C 42 | 122.8 | 2.1 | 203 | 5 | BY426544 |
| C 43 | 122.8 | 2.1 | 246 | 1 | AV244652 |
| C 44 | 122.8 | 2.1 | 278 | 4 | BM124558 |
| C 45 | 122.8 | 2.1 | 313 | 5 | BY415619 |

ALIGNMENTS

RESULT 1
AZ269661/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ269661 613 bp DNA linear GSS 26-JUL-2000
RPCI-23-444F20.TJ RPCI-23 Mus musculus genomic clone
RPCI-23-444F20, genomic survey sequence.
A2269661
A2269661 GI:9483276
GSS.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 613)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Teagaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-444F20.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@jlong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_end/mouse/bac_end_intro.html
Plate: 444 row: F column: 20
Seq primer: SP6
Class: BAC ends.

FEATURES

location/Qualifiers
1. 613
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-444F20"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
ECORI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

DEFINITION RPCI-23-117N22.TJ RPCI-23 Mus musculus genomic clone
 AZ300332.1 GI:9542117
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 496)
 REFERENCE
 AUTHORS
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
 Akinret, B., Levins, M., McGinn, S., Tsegaye, G., Geer, K., Kroll, M., de
 Jong, P. and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 OTHER_GSSs: RPCI-23-117N22.TV
 CONTACT: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@igir.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
 Plate: 117 row: N column: 22
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..496
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-117N22"
 /sex="Female"
 /lab_host="DH10B"
 /clone_lib="RPCI-23"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACe3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 8.3%; Score 473.8; DB 8; Length 496;
 Best Local Similarity 98.4%; Pred. No. 7.1e-114;
 Matches 489; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 QY CCCACTTATACCTTTCTCCCTCAGCCGCCAGATTAAACCTCTGGCCTTCCCTTCCC 3318
 DB CCCACTTATACCTTTCTCCCTCAGCCGCCAGATTAAACCTCTGGCCTTCCCTTCCC 437
 QY ACCTCCCATCAGAGTGGAGGGTTGCAGAGGAGGTTAAACCTACATGTCCCAACATC 3378
 DB ACCTCCCATCAGAGTGGAGGGTTGCAGAGGAGGTTAAACCTACATGTCCCAACATC 377
 QY ATGTGTCACGATATATGATCAGTATGTGTAGAGGCAAGAAAGGAAATCTGCAGGCTTAA 3438
 DB ATGTGTCACGATATATGATCAGTATGTGTAGAGGCAAGAAAGGAAATCTGCAGGCTTAA 318
 QY CTGGGTTAATGTGTAAGTCTGTGTGCAATGTGTGTGCTGACGTGAAACCGGGATGCG 3498
 DB CTGGGTTAATGTGTAAGTCTGTGTGCAATGTGTGTGCTGACGTGAAACCGGGATGCG 258
 QY TGTGAGCTGTTTCAGTCTGTGTCGAGGTTACACAGCTGCAGGTTTGTGTGTAATGTC 3558

Db 257 TGTGAGCTGTTTCAGTCTGTGTCGAGGTTACCAGACTGCAGGTTTGTGTGTAATGTC 198
 QY 3559 CCAAGGCAAAAGTGGGTGAATCCCTTCATGTTTAAAGAGATTGATGATGCGCTGCATC 3618
 Db 197 CCAAGGCAAAAGTGGGTGAACACCTTTCCATGTTTAAAGAGATTGATGATGCGCTGCATC 138
 QY 3619 TCAAGGACATCGGAAATAGAAATGACACTCTATATGTGTCTCTAAGCTTAAGCTAGCAAG 3678
 Db 137 TCAGGACATCGGAAATAGAAATGACACTCTATATGTGTCTCTAAGCTTAAGCTAGCAAG 78
 QY 3679 GTCTTTGGAGACACCTCTCTAGAGATGTGGCAACAGAGACTACAGACAGTATCTGTAC 3738
 Db 77 GTCTTTGGAGACACCTCTCTAGAGATGTGGCAACAGAGACTACAGACAGTATCTGCAC 18
 QY 3739 AGAGTAAGGAGAGAGAG 3755
 Db 17 AGAGTAAGGAGAGAGAG 1

RESULT 4

AZ457684/c
 LOCUS
 DEFINITION
 1M0261E12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0261E12 F, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AZ457684
 AZ457684.1 GI:10615809
 GSS.
 Mus musculus (house mouse)

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 312)
 REFERENCE
 AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished (2000)
 CONTACT: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0261 row: E column: 12
 Seq primer: GTGTGTAAGCAGCGCCAGT

Class: plasmid ends
 High quality sequence stop: 312.
 Location/Qualifiers

1..312
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0261E12"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel

FEATURES
 source

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end page: http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html
Plate: 79 row: L column: 13
Seq primer: 17
Class: BAC ends.

FEATURES

Location/Qualifiers
1..447
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-79L13"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

ORIGIN

Query Match 4.6%; Score 266.4; DB 8; Length 447;
Best Local Similarity 97.6%; Pred. No. 8.3e-59;
Matches 281; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Oy 1 GGATCTCTCAAGGTCACCAAGGGTCTCCACCACGAGTGCCTAGTCTCAATTTCAGT 60
Db 161 GGATCTCTCAAGGTCACCAAGGGTCTCCACCACGAGTGCCTAGTCTCAATTTCAGT 219
Oy 61 TTCCATGCTTGTCTCAATGCTGGCTCCCAAGAGCTAATTGGACTTTGTTTAT 120
Db 220 TTCCATGCTTGTCTCAATGCTGGCTCCCAAGAGCTAATTGGACTTTGTTTAT 279
Oy 121 TTCAAGAGGCTGATGAGGATGATCTGTGTACCCAGCTTAAGGGTCCCGTGA 180
Db 280 TTCAAGAGGCTGATGAGGATGATCTGTGTACCCAGCTTAAGGGTCCCGTGA 339
Oy 181 AGCCCTCAGACTGGAGCTTTGCAACAGCCCTTTAGGTGGAAGCAGATAAAGCAATT 240
Db 340 AGCCCTCAGACTGGAGCTTTGCAACAGCCCTTTAGGTGGAAGCAGATAAAGCAATT 399
Oy 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTCTGACCTCGG 288
Db 400 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTCTGACCTCGG 447

RESULT 7
BY000251/c
LOCUS
DEFINITION BY000251 RIKEN full-length enriched, adult male heart Mus musculus
CDNA clone 102001103 5', mRNA sequence.
ACCESSION BY000251.1 GI:26060500
VERSION BY000251.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 447)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, R., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chocho, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzle, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yang, S., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
22354683
PUBMED
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers
1..447
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="102001103"
/sex="male"
/tissue_type="heart"
/dev stage="adult"
/lab host="BNN132"
/clone lib="RIKEN full-length enriched, adult male heart"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia

FEATURES
source

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

| FEATURES | | Location/Qualifiers | |
|-----------|---|---|--|
| source | 1..390 | /organism="Mus musculus" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:10090" | |
| | | /clone="1020013103" | |
| | | /sex="male" | |
| | | /tissue_type="heart" | |
| | | /dev_stage="adult" | |
| | | /lab_host="BNI132" | |
| | | /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5]." | |
| | | GAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTTNN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5]. | |
| ORIGIN | GAGAGAGAGAGTCCAGAGCTCAATTAATTAACCCGCCCC 3'. | | |
| | cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I." | | |
| | Query Match | 2.3%; | Score 133.4; DB 5; Length 390; |
| | Best Local Similarity | 95.8%; | Pred. No. 1.7e-23; |
| | Matches 137; Conservative | 0; | Mismatches 6; Indels 0; Gaps 0; |
| | QY | 5289 | TCTCGTCAGGATCTAGATGGTCTCCAGCCTCTCTACTCTCTCTCGCTGTC 5348 |
| | Db | 10 | TCCACCCAGGATCTCGATTGGTCTCCAGCCTCTGCTCTCTCTCTCGCTGTC 69 |
| | QY | 5349 | CTCTCTCTGTCAGCTGGCCACTGTGCTGCTCTGCTGCTGCTGCTGCTGCTCA 5408 |
| | Db | 70 | CTCTCTCGTCCAGCTGGCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTCA 129 |
| | QY | 5409 | GGATTCTCTGAAAGTTAACCA 5431 |
| RESULT 11 | BY061731 | | |
| | LOCUS | BY061731 | 409 bp mRNA linear EST 06-DEC-2002 |
| | DEFINITION | BY061731 | RIKEN full-length enriched, pooled tissues, 16 days embryo, etc. Mus musculus cDNA clone 1920018N06 5', mRNA sequence. |
| | ACCESSION | BY061731 | 1 GI:26181469 |
| | VERSION | BY061731.1 | |
| | KEYWORDS | EST. | |
| | SOURCE | Mus musculus (house mouse) | |
| | ORGANISM | Mus musculus | |
| | REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 409) | |
| AUTHORS | Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chotia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gwincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., | | |

| | | | |
|----------|---------------------------|--|--|
| FEATURES | | Location/Qualifiers | |
| source | 1..409 | /organism="Mus musculus" | |
| | | /mol_type="mRNA" | |
| | | /db_xref="taxon:10090" | |
| | | /clone="1920018N06" | |
| | | /note="pooled tissues: (dev stage=16 days embryo, tissue_type=heart, sex=mix), (dev stage=16 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days embryo, tissue_type=heart, sex=mix), (dev stage=17 days embryo, tissue_type=stomach, sex=mix), (dev stage=17 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days pregnant, adult, tissue_type=amion, sex=female), (dev stage=13 days embryo, tissue_type=liver, sex=mix)" | |
| | | /days_embryo, etc." | |
| | | /clone_lib="RIKEN full-length enriched, pooled tissues, 16 days embryo, etc." | |
| | | /note="pooled tissues: (dev stage=16 days embryo, tissue_type=heart, sex=mix), (dev stage=16 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days embryo, tissue_type=heart, sex=mix), (dev stage=17 days embryo, tissue_type=stomach, sex=mix), (dev stage=17 days pregnant, adult, tissue_type=amion, sex=female), (dev stage=13 days embryo, tissue_type=liver, sex=mix)" | |
| | | /days_embryo, etc." | |
| | | /clone_lib="RIKEN full-length enriched, pooled tissues, 16 days embryo, etc." | |
| ORIGIN | Query Match | | |
| | Best Local Similarity | | |
| | Matches 137; Conservative | | |
| | Query Match | | |
| | Best Local Similarity | | |
| | Matches 137; Conservative | | |
| | Query Match | | |
| | Best Local Similarity | | |
| | Matches 137; Conservative | | |
| | Query Match | | |

Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Alizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers

1..409

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone_lib="RIKEN full-length enriched, pooled tissues, 16 days embryo, etc."

/note="pooled tissues: (dev stage=16 days embryo, tissue_type=heart, sex=mix), (dev stage=16 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days embryo, tissue_type=heart, sex=mix), (dev stage=17 days embryo, tissue_type=stomach, sex=mix), (dev stage=17 days pregnant, adult, tissue_type=amion, sex=female), (dev stage=13 days embryo, tissue_type=liver, sex=mix)"

Query Match

2.3%;

Score 133.4; DB 5; Length 409;

Best Local Similarity

95.8%;

Pred. No. 1.7e-23;

Matches 137; Conservative

0;

Mismatches 6;

Indels 0;

Gaps 0;


```

/mol_type="mrna"
/db_xref="taxon:10090"
/clone="IMAGE:6755570"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_156"
/notes="Organ: Heart; TCCD (DMSO vehicle) treated 48 hours
IP injection; Vector: pDONR201; Site 1: attB2; Site 2:
attP; cDNA made by oligo-dT with attB2 site and
directionally cloned. Priming sequence:
5'-TTTCTGTGAGCCGCCACCACTTGTACAGAAAGCTGGGTTTTTTTTTTT
TTTT-3'. Full-length enriched library was constructed
using the Genekacer kit by Invitrogen, library
amplification 16 cycles. Library constructed by Mark
Bittinger in the Bradfield laboratory (McArdle Laboratory
for Cancer Research, University of Wisconsin). Note: this
is a NIH_MGC Library."

```

ORIGIN

```

Query Match      2.3%; Score 133.4; DB 6; Length 783;
Best Local Similarity 95.8%; Pred. No. 2.2e-23;
Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5289 TCCTGGTCAGGATCTCTAGATTGGTCTCCAGCCTCTGCTACTCTTCTTCTGCTGTC 5348
Db 51  TCCAAACCCAGGATCTCTGATTGGTCTCCAGCCTCTGCTACTCTTCTTCTGCTGTC 110

Qy 5349 CTCTCTCTGTCCAGTGGCCACTGTGGTGGCTGCTCCAGCTGTGGTCCACATTTCTCA 5408
Db 111 CTCTCTCGTCCAGTGGCCACTGTGGTGGCTGCTCCAGCTGTGGTCCACATTTCTCA 170

Qy 5409 GGATTCTCTGAAAAGTTAAACCAG 5431
Db 171 GGATTCTCTGAAAAGTTAAACCAG 193

```

Search completed: April 25, 2005, 14:31:21
Job time : 12067.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 02:43:47 ; Search time 636.574 Seconds
(without alignments)
14741.485 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatccgcgaaggtcacaca.....ccataaggtttgagtcgac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 1279.8 | 22.3 | 1679 | 2 | US-08-365-486A-11 |
| 2 | 1279.8 | 22.3 | 1679 | 3 | US-08-880-342-11 |
| 3 | 312.6 | 5.5 | 10728 | 4 | US-09-376-774-5 |
| 4 | 311.8 | 5.4 | 520 | 1 | US-08-485-971-7 |
| 5 | 311.8 | 5.4 | 520 | 1 | US-08-275-876-7 |
| 6 | 311.8 | 5.4 | 520 | 1 | US-08-383-754-7 |
| 7 | 311.8 | 5.4 | 520 | 1 | US-08-485-978-7 |
| 8 | 311.8 | 5.4 | 520 | 2 | US-08-486-814-7 |
| 9 | 311.8 | 5.4 | 520 | 2 | US-08-487-472-7 |
| 10 | 311.8 | 5.4 | 520 | 3 | US-08-485-740-7 |
| 11 | 311.8 | 5.4 | 520 | 3 | US-09-162-184-7 |
| 12 | 311.8 | 5.4 | 520 | 3 | US-09-161-902-7 |
| 13 | 311.8 | 5.4 | 520 | 3 | US-09-489-777A-7 |
| 14 | 311.8 | 5.4 | 520 | 5 | PCT-US95-08179-7 |
| 15 | 310.4 | 5.4 | 450 | 1 | US-08-076-726-13 |
| 16 | 310.4 | 5.4 | 450 | 1 | US-08-485-971-9 |
| 17 | 310.4 | 5.4 | 450 | 1 | US-08-260-452-6 |
| 18 | 310.4 | 5.4 | 450 | 1 | US-08-275-876-9 |
| 19 | 310.4 | 5.4 | 450 | 1 | US-08-383-754-9 |
| 20 | 310.4 | 5.4 | 450 | 1 | US-08-485-978-9 |
| 21 | 310.4 | 5.4 | 450 | 2 | US-08-481-970-6 |
| 22 | 310.4 | 5.4 | 450 | 2 | US-08-486-814-9 |
| 23 | 310.4 | 5.4 | 450 | 2 | US-08-487-472-9 |
| 24 | 310.4 | 5.4 | 450 | 2 | US-08-897-719-6 |
| 25 | 310.4 | 5.4 | 450 | 3 | US-08-485-740-9 |
| 26 | 310.4 | 5.4 | 450 | 3 | US-09-162-184-9 |
| 27 | 310.4 | 5.4 | 450 | 3 | US-09-161-902-9 |

c 28 310.4 5.4 450 3 US-09-163-269-6 Sequence 6, Appli
c 29 310.4 5.4 450 3 US-09-489-777A-9 Sequence 9, Appli
c 30 310.4 5.4 450 4 US-09-281-674-6 Sequence 6, Appli
c 31 310.4 5.4 450 5 PCT-US95-08179-9 Sequence 9, Appli
32 310.4 5.4 450 1 US-08-076-726-12 Sequence 12, Appli
33 309.4 5.4 450 1 US-08-485-971-8 Sequence 8, Appli
34 309.4 5.4 450 1 US-08-260-452-5 Sequence 5, Appli
35 309.4 5.4 450 1 US-08-275-876-8 Sequence 8, Appli
36 309.4 5.4 450 1 US-08-383-754-8 Sequence 8, Appli
37 309.4 5.4 450 1 US-08-485-978-8 Sequence 8, Appli
38 309.4 5.4 450 2 US-08-481-970-5 Sequence 5, Appli
39 309.4 5.4 450 2 US-08-486-814-8 Sequence 8, Appli
40 309.4 5.4 450 2 US-08-487-472-8 Sequence 8, Appli
41 309.4 5.4 450 2 US-08-897-719-5 Sequence 5, Appli
42 309.4 5.4 450 3 US-08-485-740-8 Sequence 8, Appli
43 309.4 5.4 450 3 US-09-162-184-8 Sequence 8, Appli
44 309.4 5.4 450 3 US-09-161-902-8 Sequence 8, Appli
45 309.4 5.4 450 3 US-09-163-269-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-365-486A-11
; Sequence 11, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 82555-0018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1679 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Mouse alpha MHC promoter fragment
US-08-365-486A-11

Query Match 22.3%; Score 1279.8; DB 2; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

QY 5267 CCCCTGGCTTCTGCTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCAGCCTCTG 5326
Db 1209 CCCCTGGCTTCTGCTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCAGCCTCTG 1268
QY 5327 CTACTCTCTCTCTGCTGCTTCTCTCTCTGTCAGCTGCGCACTGTGGTCCCTCGTTC 5386
Db 1269 CTACTCTCTCTCTGCTGCTTCTCTCTCTGTCAGCTGCGCACTGTGGTCCCTCGTTC 1328
QY 5387 CAGCTGTGTCACATCTCTCAGGATCTCTGAAAAGTTAACCCAGTCAGATGTTTCCC 5446
Db 1329 CAGCTGTGTCACATCTCTCAGGATCTCTGAAAAGTTAACCCAGTCAGATGTTTCCC 1388
QY 5447 CTGTAGACAGCAGATCAGATCTCTCCGGAAGTCAGGCTTCAGGCCCTCTCTTCTCTGC 5506
Db 1389 CTGTAGACAGCAGATCAGATCTCTCCGGAAGTCAGGCTTCAGGCCCTCTCTTCTCTGC 1448
QY 5507 CCAGCTGCGCGCACTCTTAGCAAACTTCAGGCACTTACCCACATAGACCTCTGACA 5566
Db 1449 CCAGCTGCGCGCACTCTTAGCAAACTTCAGGCACTTACCCACATAGACCTCTGACA 1508
QY 5567 GAGAGCAGGCACTTACATGAGTCCTGCTGGGAGGCCATGAGCTACGGTGTAAAGA 5626
Db 1509 GAGAGCAGGCACTTACATGAGTCCTGCTGGGAGGCCATGAGCTACGGTGTAAAGA 1568
QY 5627 GCGAGGGAAGTGGTGTAGAAAAGTCAGGACTTCACATAGAGCCTAGCCACACAG 5686
Db 1569 GCGAGGGAAGTGGTGTAGAAAAGTCAGGACTTCACATAGAGCCTAGCCACACAG 1628
QY 5687 AAATGACAGACAGATCCCTCTCTATCTCCGCCATAGAGTTGAGT 5731
Db 1629 AAATGACAGACAGATCCCTCTCTATCTCCGCCATAGAGTTGAGT 1673

RESULT 3

US-09-376-774-5

; Sequence 5, Application US/09376774

; Patent No. 6759236

; GENERAL INFORMATION:

; APPLICANT: Fung, Yuen Kai

; APPLICANT: Gomer, Charles

; APPLICANT: T'Ang, Anne

; TITLE OF INVENTION: Methods To Enhance And Confine Expression

; FILE REFERENCE: Of Genes

; CURRENT APPLICATION NUMBER: US/09/376,774

; CURRENT FILING DATE: 2003-03-21

; PRIOR APPLICATION NUMBER: 60/096,947

; PRIOR FILING DATE: 1998-08-18

; NUMBER OF SEQ ID NOS: 5

; SEQ ID NO 5

; LENGTH: 10728

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: recombinant vector pDATH-TNP?

US-09-376-774-5

. Query Match . 5.5%; Score 312.6; DB 4; Length 10728;

Best Local Similarity 97.2%; Pred. No. 7.7e-82;

Matches 318; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4265 AACGACAGGAGGAACTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 4324
Db 212 ATCCGACAGAAAGCTTGTGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 271
QY 4325 CGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGT 4384
Db 272 CGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGT 331
QY 4385 TCAGTGATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4444
Db 332 TCAGTGATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 391

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-485-971-7

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAATCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCC 4336
Db 63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCC 122
QY 4337 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAG 182
QY 4397 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 242
QY 4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
Db 243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGT 4576
Db 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
Db 363 CGAGCTCGGTACCG 377

RESULT 5

US-08-275-876-7
Sequence 7, Application US/08275876
Patent No. 5654168
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/275,876
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 01-JULY-94
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-275-876-7

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAATCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCC 4336
Db 63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCC 122
QY 4337 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAG 182
QY 4397 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 242
QY 4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
Db 243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGT 4576
Db 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
Db 363 CGAGCTCGGTACCG 377

RESULT 6

US-08-383-754-7
Sequence 7, Application US/08383754
Patent No. 5789156
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,754
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-94
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-94
CLASSIFICATION: 436

[illegible]

| | | | |
|----|------|--|------|
| Qy | 4337 | TCCCTATCAGTGTAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCCTATCAGTGTATAGAG | 4396 |
| Db | 123 | TCCCTATCAGTGTAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCCTATCAGTGTATAGAG | 182 |
| Qy | 4397 | AAAAGTGAAGTCGAGTTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAAAGTCGAGTT | 4456 |
| Db | 183 | AAAAGTGAAGTCGAGTTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAAAGTCGAGTT | 242 |
| Qy | 4457 | TACCACTCCCCTATCAGTGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCCTATCAGTG | 4516 |
| Db | 243 | TACCACTCCCCTATCAGTGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCCTATCAGTG | 302 |
| Qy | 4517 | ATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCCTATCAGTGTATAGAGAAAAGTGAAAAGT | 4576 |
| Db | 303 | ATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCCTATCAGTGTATAGAGAAAAGTGAAAAGT | 362 |
| Qy | 4577 | CGAGTCGGTACCAG | 4591 |
| Db | 363 | CGAGTCGGTACCG | 377 |

```

RESULT 7
US-08-485-978-7
; Sequence 7, Application US/08485978
; Patent No. 5814618
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; APPLICANT: Gossen, Manfred
; TITLE OF INVENTION: Methods for Regulating Gene Expression
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

RESULT 8
US-08-486-814-7
; Sequence 7, Application US/08486814
; Patent No. 5866755
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; APPLICANT: Gossen, Manfred
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline-
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,814
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/383,754
; FILING DATE: 03-FEB-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,876
; FILING DATE: 15-JULY-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/270,637
; FILING DATE: 01-JULY-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,452
; FILING DATE: 14-JUNE-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/076,726
; FILING DATE: 14-JUNE-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-009CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-486-814-7

Query Match 5.4%; Score 311.8; DB 2; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4277 GGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCAC 4336
Db 63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCAC 122

Qy 4337 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182
Qy 4397 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTT 242
Qy 4457 TACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
Db 243 TACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
Qy 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTC 4576
Db 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTC 362
Qy 4577 CGAGCTCGGTACCAG 4591
Db 363 CGAGCTCGGTACCG 377

RESULT 9

US-08-487-472-7
; Sequence 7, Application US/08487472
; Patent No. 5912411
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; APPLICANT: Gossen, Manfred
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,472
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/383,754
; FILING DATE: 03-FEB-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,876
; FILING DATE: 15-JULY-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/270,637
; FILING DATE: 01-JULY-94
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,452
; FILING DATE: 14-JUNE-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/076,327
; FILING DATE: 14-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/076,726
; FILING DATE: 14-JUNE-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.

[illegible]

Db 63 GGATCTCAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCAC 122
Qy 4337 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182
Qy 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTT 242
Qy 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTG 4516
Db 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTG 302
Qy 4517 ATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGT 4576
Db 303 ATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGT 362
Qy 4577 CGAGCTCGGTACCAG 4591
Db 363 CGAGCTCGGTACCAG 377

RESULT 13

US-09-489-777A-7
; Sequence 7, Application US/09489777A
; Patent No. 6271348
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; Gossen, Manfred
; TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
; Inhibitor Fusion Proteins
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09489,777A
; FILING DATE: 24-Jan-2000
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/162,184
; FILING DATE: 28-SEP-1998
; APPLICATION NUMBER: US 08/485,978
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/383,754
; FILING DATE: 03-FEB-1995
; APPLICATION NUMBER: US 08/275,876
; FILING DATE: 15-JULY-1994
; APPLICATION NUMBER: US 08/270,637
; FILING DATE: 01-JULY-1994
; APPLICATION NUMBER: US 08/260,452
; FILING DATE: 14-JUNE-1994
; APPLICATION NUMBER: US 08/076,327
; FILING DATE: 14-JUNE-1993
; APPLICATION NUMBER: US 08/076,726
; FILING DATE: 14-JUNE-1993

ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-489-777A-7

Query Match 5.4%; Score 311.8; DB 3; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4277 GGAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCAC 4336
Db 63 GGATCTCAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCAC 122
Qy 4337 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182
Qy 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTT 242
Qy 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTG 4516
Db 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTG 302
Qy 4517 ATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGT 4576
Db 303 ATAGAGAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAAGT 362
Qy 4577 CGAGCTCGGTACCAG 4591
Db 363 CGAGCTCGGTACCAG 377

RESULT 14

PCT-US95-08179-7
; Sequence 7, Application PC/TUS9508179
; GENERAL INFORMATION:
; APPLICANT: Tetracycline-Regulated Transcriptional
; TITLE OF INVENTION: Modulators
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08179
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: To be assigned
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/383,754
; FILING DATE: 03-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-94
CLASSIFICATION: B1-009C2PC
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION: B1-009C2PC
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C2PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-08179-7

Query Match 5.4%; Score 311.8; DB 5; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4277 GGAATCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
Db 63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCAC 122
Qy 4337 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAG 182
Qy 4397 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTT 242
Qy 4457 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 4516
Db 243 TACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTG 302
Qy 4517 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGT 4576
Db 303 ATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGT 362
Qy 4577 CGAGTCGCTACCG 4591
Db 363 CGAGTCGCTACCG 377

RESULT 15
US-08-076-726-13/c
Sequence 13, Application US/08076726
Patent No. 5464758
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
TITLE OF INVENTION: Tight Control of Gene Expression in
TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox
CITY: Washington
STATE: District of Columbia
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,726
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2490001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-076-726-13

Query Match 5.4%; Score 310.4; DB 1; Length 450;
Best Local Similarity 99.7%; Pred. No. 4.1e-82;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4280 ACTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCC 4339
Db 325 ACTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCC 266
Qy 4340 CTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAA 4399
Db 265 CTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAA 206
Qy 4400 AGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTAC 4459
Db 205 AGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTAC 146
Qy 4460 CACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATA 4519
Db 145 CACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATA 86
Qy 4520 GAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGA 4579
Db 85 GAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAGTGAAGTCGA 26
Qy 4580 GCTCGGTACCG 4591
Db 25 GCTCGGTACCG 14

Search completed: April 25, 2005, 14:46:02
Job time : 643.574 secs

THIS PAGE BLANK (USPTO)

Montanari, D.
10/613728
Seq IDs 1# 5-8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 25, 2005, 00:13:02 ; Search time 16284 seconds
(without alignments)
17065.210 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatcctgaggtcacaca.....ccataagagttgagtcgac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query | ID | Description |
|------------|--------|-------|--------|----|---------------------|
| 1 | 5045 | 88.0 | 5443 | 6 | BD140880 An animal |
| 2 | 5045 | 88.0 | 5443 | 10 | U71441 Mus musculus |
| 3 | 4804.8 | 83.8 | 235759 | 2 | AC099577 Mus muscu |
| 4 | 2343 | 40.9 | 185702 | 2 | AC130940 Rattus no |
| 5 | 2343 | 40.9 | 210784 | 2 | AC119293 Rattus no |
| 6 | 2320.2 | 40.5 | 5487 | 10 | AV191158 Rattus no |
| 7 | 1655.4 | 28.9 | 263901 | 2 | AC115371 Rattus no |
| 8 | 1444.8 | 25.2 | 32415 | 10 | LI5351 Mesocricetu |
| 9 | 1279.8 | 22.3 | 1679 | 6 | AR054006 Sequence |
| 10 | 1279.8 | 22.3 | 1679 | 6 | AR146182 Sequence |
| 11 | 1279.8 | 22.3 | 1679 | 6 | AX546496 Sequence |
| 12 | 1279.8 | 22.3 | 2594 | 10 | M62404 Mus musculu |
| 13 | 558.2 | 9.7 | 157310 | 9 | AL132855 Human chr |
| 14 | 558.2 | 9.7 | 196292 | 9 | AL049829 Human chr |
| 15 | 494.8 | 8.6 | 31462 | 9 | Z20656 Homo sapien |
| 16 | 462.6 | 8.1 | 516 | 11 | BV093112 RPAMMSEQ0 |
| 17 | 462.4 | 8.1 | 542 | 11 | BV163870 RPAMMSEQ0 |
| 18 | 438.4 | 7.6 | 452 | 11 | BV093084 RPAMMSEQ0 |
| 19 | 424.4 | 7.4 | 641 | 10 | K01463 Rat cardiac |

| | | | | | | |
|----|-------|-----|-------|----|----------|--------------------|
| 20 | 407.6 | 7.1 | 6020 | 4 | AF192305 | AF192305 Oryctolag |
| 21 | 396.4 | 6.9 | 541 | 11 | BV159459 | BV159459 RPAMMSEQ0 |
| 22 | 395.4 | 6.9 | 538 | 11 | BV093111 | BV093111 RPAMMSEQ0 |
| 23 | 378 | 6.6 | 580 | 11 | BV159460 | BV159460 RPAMMSEQ0 |
| 24 | 377 | 6.6 | 581 | 11 | BV093113 | BV093113 RPAMMSEQ0 |
| 25 | 343.6 | 6.0 | 2450 | 6 | BD096863 | BD096863 Agent for |
| 26 | 312.6 | 5.5 | 10728 | 6 | AR562497 | AR562497 Sequence |
| 27 | 311.8 | 5.4 | 520 | 6 | AR020198 | AR020198 Sequence |
| 28 | 311.8 | 5.4 | 520 | 6 | AR032150 | AR032150 Sequence |
| 29 | 311.8 | 5.4 | 520 | 6 | AR043815 | AR043815 Sequence |
| 30 | 311.8 | 5.4 | 520 | 6 | AR072127 | AR072127 Sequence |
| 31 | 311.8 | 5.4 | 520 | 6 | AR095981 | AR095981 Sequence |
| 32 | 311.8 | 5.4 | 520 | 6 | AR136641 | AR136641 Sequence |
| 33 | 311.8 | 5.4 | 520 | 6 | AR157176 | AR157176 Sequence |
| 34 | 311.8 | 5.4 | 520 | 6 | AR164125 | AR164125 Sequence |
| 35 | 311.8 | 5.4 | 520 | 6 | I32811 | I32811 Sequence 7 |
| 36 | 311.8 | 5.4 | 520 | 6 | I59625 | I59625 Sequence 7 |
| 37 | 311.8 | 5.4 | 520 | 6 | AX127248 | AX127248 Sequence |
| 38 | 310.8 | 5.4 | 5961 | 12 | AF326320 | AF326320 Shuttle v |
| 39 | 310.4 | 5.4 | 450 | 6 | AR020200 | AR020200 Sequence |
| 40 | 310.4 | 5.4 | 450 | 6 | AR029415 | AR029415 Sequence |
| 41 | 310.4 | 5.4 | 450 | 6 | AR032152 | AR032152 Sequence |
| 42 | 310.4 | 5.4 | 450 | 6 | AR043817 | AR043817 Sequence |
| 43 | 310.4 | 5.4 | 450 | 6 | AR072129 | AR072129 Sequence |
| 44 | 310.4 | 5.4 | 450 | 6 | AR095983 | AR095983 Sequence |
| 45 | 310.4 | 5.4 | 450 | 6 | AR136643 | AR136643 Sequence |

ALIGNMENTS

| | | | | | | |
|-----------------------|---|--|------------|-----|--------|-----------------|
| RESULT 1 | BD140880 | BD140880 | 5443 bp | DNA | linear | PAT 18-SEP-2002 |
| LOCUS | BD140880 | An animal into which p300 gene is introduced. | | | | |
| DEFINITION | BD140880 | | | | | |
| ACCESSION | BD140880.1 | GI:23235825 | | | | |
| VERSION | WO 0205633-A/2. | | | | | |
| KEYWORDS | Mus musculus (house mouse) | | | | | |
| SOURCE | Mus musculus | | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus | | | | | |
| REFERENCE | 1 (bases 1 to 5443) | | | | | |
| AUTHORS | Hasegawa, K., Kawase, Y. and Suzuki, H. | | | | | |
| TITLE | An animal into which p300 gene is introduced | | | | | |
| JOURNAL | Patent: WO 0205633-A 2 24-JAN-2002; | | | | | |
| | KOJI HASEGAWA, CHUGAI PHARMACEUTICAL CO LTD, YOSUKE KAWASE, HIROSHI SUZUKI | | | | | |
| COMMENT | OS Mus musculus (mouse) | | | | | |
| | PN WO 0205633-A/2 | | | | | |
| | PD 24-JAN-2002 | | | | | |
| | PF 13-JUL-2001 WO 2001JP006086 | | | | | |
| | PR 14-JUL-2000 JP 00P 215143 | | | | | |
| | PI KOJI HASEGAWA, YOSUKE KAWASE, HIROSHI SUZUKI | | | | | |
| | PC A01K67/027, C12N15/09, C12N15/12, C12Q1/68, A61K45/00, A61P9/04, PC GOIN33/50, | | | | | |
| | GOIN33/15 | | | | | |
| PC | GOIN33/15 | | | | | |
| CC | An animal into which p300 gene is introduced | | | | | |
| FF | Key Location/Qualifiers | | | | | |
| FT | promoter (1)..(5443). | | | | | |
| FEATURES | Location/Qualifiers | | | | | |
| source | 1..5443 | | | | | |
| | /organism="Mus musculus" | | | | | |
| | /mol_type="genomic DNA" | | | | | |
| | /db_xref="taxon:10090" | | | | | |
| ORIGIN | | | | | | |
| Query Match | 88.0%; | Score | 5045; | DB | 6; | Length 5443; |
| Best Local Similarity | 93.9%; | Pred. | No. 0; | | | |
| Matches 5383; | Conservative | 0; | Mismatches | 60; | Indels | 292; |
| Gaps | 1; | | | | | |
| Qy. | 1 | GGATCCTGCAAGGTGACAAAGGCTTCACCCACGAGGTGCCTAGCTCAATTCAGT | 60 | | | |

QY 2221 AAGGAAGAGGCGAAACAGGCCACACAGAGGCGAGCCAGAACTGAGTTAACTC 2280
DB 2221 AAGGAAGAGGCGAAACAGGCCACACAGAGGCGAGCCAGAACTGAGTTAACTC 2280
QY 2281 CTTCTCTGTTGTCATCTTCATAGGAGGAGTGGGAACTCTGTGCAACCACTCCCCATGA 2340
DB 2281 CTTCTCTGTTGTCATCTTCATAGGAGGAGTGGGAACTCTGTGCAACCACTCCCCATGA 2340
QY 2341 GCCCCACTACCATACCAAGTTGGCCTGAGTGGCAATCTAGGTTCCCTGAGGACAGAG 2400
DB 2341 GCCCCACTACCATACCAAGTTGGCCTGAGTGGCAATCTAGGTTCCCTGAGGACAGAG 2400
QY 2401 CTTGGCCTTTGTCCTTTGACCTGACCCAAAGTGCACCAATGTTCTCAGTACCTTTGAAT 2460
DB 2401 CTTGGCCTTTGTCCTTTGACCTGACCCAAAGTGCACCAATGTTCTCAGTACCTTTGAAT 2460
QY 2461 GCCCTCAAGAGCTTGAGAAACAGGAGTGCATATATTAGGCCATGGGCTAACCTGAGGT 2520
DB 2461 GCCCTCAAGAGCTTGAGAAACAGGAGTGCATATATTAGGCCATGGGCTAACCTGAGGT 2520
QY 2521 TGCACACAGGAGCCTCAAGTACCTCCAGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
DB 2521 TGCACACAGGAGCCTCAAGTACCTCCAGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
QY 2581 CCAGAGAGCAACATTTGGCATAGGTGGCTGCAGAAATGGGAATGCAAGGTTGAATCAGGTC 2640
DB 2581 CCAGAGAGCAACATTTGGCATAGGTGGCTGCAGAAATGGGAATGCAAGGTTGAATCAGGTC 2640
QY 2641 CTTTCAAGAACTACTGCATCAAGACCTAAGACCTCCAGGAGAGGGGTATGCTCTGCCC 2700
DB 2641 CTTTCAAGAACTACTGCATCAAGACCTAAGACCTCCAGGAGAGGGGTATGCTCTGCCC 2700
QY 2701 CCACCCACCATAAAGGAGTGAATCTCTAGGGGCTGGGACCTTGGGGAGACACAC 2760
DB 2701 CCACCCACCATAAAGGAGTGAATCTCTAGGGGCTGGGACCTTGGGGAGACACAC 2760
QY 2761 ATTACTGAGGTGCTGAGCCAGAGAACTGACCGCCCTGTGTCTGCTGCCACTCCACAC 2820
DB 2761 ATTACTGAGGTGCTGAGCCAGAGAACTGACCGCCCTGTGTCTGCTGCCACTCCACAC 2820
QY 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGAGCTGGTAGCAGGAGAGTGT 2880
DB 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGAGCTGGTAGCAGGAGAGTGT 2880
QY 2881 TCTGGGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGAGTGTGGCTTTGCTCTGTA 2940
DB 2881 TCTGGGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGAGTGTGGCTTTGCTCTGTA 2940
QY 2941 CACAATGTCTACTTGTATACAGGCATGACCTGCTAAGACCCCAACATCTACGACCTC 3000
DB 2941 CACAATGTCTACTTGTATACAGGCATGACCTGCTAAGACCCCAACATCTACGACCTC 3000
QY 3001 TGAAGAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGTGTTG 3060
DB 3001 TGAAGAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGTGTTG 3060
QY 3061 CCAAAAGAGGAGCATGAGTGTGATATAGGCCCCCAGAGGCGTTAGAGAGGCACTTG 3120
DB 3061 CCAAAAGAGGAGCATGAGTGTGATATAGGCCCCCAGAGGCGTTAGAGAGGCACTTG 3120
QY 3121 GGAAGGGTCAGTCTGCAGAGCCCTATCCATGGATCTGGAGCCTGGGGCCCACTGGTG 3180
DB 3121 GGAAGGGTCAGTCTGCAGAGCCCTATCCATGGATCTGGAGCCTGGGGCCCACTGGTG 3180
QY 3181 TAAATCTCTGGGCTGCCAGGATTCAGAGCAGACCTGCATCTCTGCGAGCTGGGGA 3240
DB 3181 TAAATCTCTGGGCTGCCAGGATTCAGAGCAGACCTGCATCTCTGCGAGCTGGGGA 3240
QY 3241 GCGGAAGGAGCAACCCCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAACACCT 3300
DB 3241 GCGGAAGGAGCAACCCCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAACACCT 3300
QY 3301 CTGGCCTTCCCTCCACCTCCCATCAGGAGTGGAGGTTGCAGAGGAGGGTAAAAA 3360

DB 3301 CTGGCCTTCCCTCCACCTCCCATCAGGAGTGGAGGTTGCAGAGGAGGGTAAAAA 3360
QY 3361 CCTACATGTCCAAACATCATGTTGTCACATATATGATCATGATGTGTAGAGCAAGAAA 3420
DB 3361 CCTACATGTCCAAACATCATGTTGTCACATATATGATCATGATGTGTAGAGCAAGAAA 3420
QY 3421 GGAATCTGCGAGCTTAACTGGGTTAACTGTGTAAGAGTCTGTGTCATGTGTGTCTG 3480
DB 3421 GGAATCTGCGAGCTTAACTGGGTTAACTGTGTAAGAGTCTGTGTCATGTGTGTCTG 3480
QY 3481 ACTGAAAACGGGCATGGCTGTGTCAGTGTCTGTGCGTGAGGTTACAGACTGCA 3540
DB 3481 ACTGAAAACGGGCATGGCTGTGTCAGTGTCTGTGCGTGAGGTTACAGACTGCA 3540
QY 3541 GGTGTTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCATGTTTTAAAGAGAT 3600
DB 3541 GGTGTTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCATGTTTTAAAGAGAT 3600
QY 3601 TGGATGATGGCTGTCATCTCAAGGACCATGGAATAATAGAAATGACACTCTATATGTCT 3660
DB 3601 TGGATGATGGCTGTCATCTCAAGGACCATGGAATAATAGAAATGACACTCTATATGTCT 3660
QY 3661 CTAAGCTAAGGTAGCAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
DB 3661 CTAAGCTAAGGTAGCAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
QY 3721 TACAGACAGTATCTGTACAGATAAGGAGAGAGAGGGGTGTAGAAATCTCTTACTA 3780
DB 3721 TACAGACAGTATCTGTACAGATAAGGAGAGAGAGGGGTGTAGAAATCTCTTACTA 3780
QY 3781 TCAGAGGGAATCTGAGTGTGTCACCTGCAAGTGTGATCTCTCCCTAGACATCATGCTT 3840
DB 3781 TCAGAGGGAATCTGAGTGTGTCACCTGCAAGTGTGATCTCTCCCTAGACATCATGCTT 3840
QY 3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGGCTCTGAGAGTAGGAGGTCCCTCTCA 3900
DB 3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGGCTCTGAGAGTAGGAGGTCCCTCTCA 3900
QY 3901 GCTTGAAGCTATGAGATAGCCAGGGTTGAAAGGGGGAAGGAGAGCCTGGGATGGGAGC 3960
DB 3901 GCTTGAAGCTATGAGATAGCCAGGGTTGAAAGGGGGAAGGAGAGCCTGGGATGGGAGC 3960
QY 3961 TTGTGTGTGGAGGAGGAGCAGATATTAAGCTTGAAGAGAAAGTGAACCTTACCCAG 4020
DB 3961 TTGTGTGTGGAGGAGGAGCAGATATTAAGCTTGAAGAGAAAGTGAACCTTACCCAG 4020
QY 4021 TTGTTCACCTCACTCAGATTAAAAATAAAGTGAAGGCTTGGGTAGGGAGGTG 4080
DB 4021 TTGTTCACCTCACTCAGATTAAAAATAAAGTGAAGGCTTGGGTAGGGAGGTG 4080
QY 4081 GTGTGAGAGCT 4140
DB 4081 GTGTGAGAGCT 4140
QY 4141 GCCCAAGGACTTAAAAAGGCGCATCGAGCAGAGGGGAGGCAACAGACCTTTTCATGG 4200
DB 4141 GCCCAAGGACTTAAAAAGGCGCATCGAGCAGAGGGGAGGCAACAGACCTTTTCATGG 4200
QY 4201 GCAGACCTTGGGGCCGTAGTGTGATGCAAGAACTCGCAATTCGATACCTTCTTTC 4260
DB 4201 GCAGACCTTGGGGCCGTAGTGTGATGCAAGAACTCGCAATTCGATACCTTCTTTC 4260
QY 4261 TTCTAACCGAGAGGAGGAACTCGAGTTTACCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4320
DB 4261 GGAGCCAGGACAGAGGAGGAGTGGGAGGAG----- 4291
QY 4321 AAGTCGAGTTTACCCT 4380
DB 4292 ----- 4291
QY 4381 CCTATCAGTATAGAGAAAGTGAAGTGCAGTGTTCACCACTCTCTCTCTCTCTCTCTCTCT 4440

QY 301 AGGGTGGGAGGTGGGGCTTGGGAAGAAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
Db 301 AGGGTGGGAGGTGGGGCTTGGGAAGAAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
QY 361 GGGCCCTGTGAAGTTGGAGGCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420
Db 361 GGGCCCTGTGAAGTTGGAGGCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420
QY 421 AAAACATAGCAAGAGTATACCTCTCTTGTGACTTCCCGAGGCCAGATACCTGTACGGTT 480
Db 421 AAAACATAGCAAGAGTATACCTCTCTTGTGACTTCCCGAGGCCAGATACCTGTACGGTT 480
QY 481 GAAACAGAGTTTAGGAAGCCCTCTGAACCTCACTCTGAAGCTCTCAACCAAGCA 540
Db 481 GAAACAGAGTTTAGGAAGCCCTCTGAACCTCACTCTGAAGCTCTCAACCAAGCA 540
QY 541 AGCACTAGGTCCTAGTGTAGTATCTTACGCTGATATATATGCAAGCTGGGCCAC 600
Db 541 AGCACTAGGTCCTAGTGTAGTATCTTACGCTGATATATATGCAAGCTGGGCCAC 600
QY 601 AGAAGTCTGGGGTGTAGGAACCTGACCACTGACTTTTCACTCGGCAAGGTATGACCCCC 660
Db 601 AGAAGTCTGGGGTGTAGGAACCTGACCACTGACTTTTCACTCGGCAAGGTATGACCCCC 660
QY 661 TCAGCAGATGTAGTATGTCCTTAGATCCCATCCAGGCAAGTCTCTAAGAGGACATG 720
Db 661 TCAGCAGATGTAGTATGTCCTTAGATCCCATCCAGGCAAGTCTCTAAGAGGACATG 720
QY 721 GGATGAGAGATGTAGTATGTGGCATTCCAACACACAGCTATCCACAGTGTCCCTTGCCCC 780
Db 721 GGATGAGAGATGTAGTATGTGGCATTCCAACACACAGCTATCCACAGTGTCCCTTGCCCC 780
QY 781 TTCCCACTTAGCCAGGAGCAGTAACTTAGCTATCTTTCTTCCCTCCCATCTCCCCAG 840
Db 781 TTCCCACTTAGCCAGGAGCAGTAACTTAGCTATCTTTCTTCCCTCCCATCTCCCCAG 840
QY 841 GACACACCCCTGGTCTGCAGTATTCATTTCTTCTTCACTGCTCCCTCTGCTACCTTCCAT 900
Db 841 GACACACCCCTGGTCTGCAGTATTCATTTCTTCTTCACTGCTCCCTCTGCTACCTTCCAT 900
QY 901 TTGCAAGGCTTTTGACCTCTGCAGCTGTGGAGATAGAGTTTGGCCCTTAGGTGTGGCAA 960
Db 901 TTGCAAGGCTTTTGACCTCTGCAGCTGTGGAGATAGAGTTTGGCCCTTAGGTGTGGCAA 960
QY 961 GCATCTCAAGAAAGCAGACAAACAGGGGACAGATTTTGAAGGATCAGAACTAAA 1020
Db 961 GCATCTCAAGAAAGCAGACAAACAGGGGACAGATTTTGAAGGATCAGAACTAAA 1020
QY 1021 TCACCTGGCGGGCTGGGGGTAGAAAAGAGTGAAGTGAAGTCCGCTCCAGCTAAGCCAGC 1080
Db 1021 TCACCTGGCGGGCTGGGGGTAGAAAAGAGTGAAGTGAAGTCCGCTCCAGCTAAGCCAGC 1080
QY 1081 TAGTCCCGAGATATCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
Db 1081 TAGTCCCGAGATATCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
QY 1141 TGAAGAACAATGGGATTTGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
Db 1141 TGAAGAACAATGGGATTTGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
QY 1201 ACACTGCTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTTCTGACCACTGAGCTAG 1260
Db 1201 ACACTGCTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTTCTGACCACTGAGCTAG 1260
QY 1261 GCAACACAGAGTGGGCCCTGTGCTGAGATGAAGTGTGGTTTACCAATAGCAAAAACAG 1320
Db 1261 GCAACACAGAGTGGGCCCTGTGCTGAGGATGAAGTGTGGTTTACCAATAGCAAAAACAG 1320
QY 1321 CAGGGAGGGAGAACAGAGAACCAATAAGAGGAGAGGAAAGGCCAGTCAATCAGA 1380
Db 1321 CAGGGAGGGAGAACAGAGAACCAATAAGAGGAGAGGAAAGGCCAGTCAATCAGA 1380
QY 1381 TCCAGTCAGAGATGGGAAGCCACACACAGCTTGACGAGGAAACAGAAAAGGGAG 1440

Db 1381 TCCAGTCAGAGAGATGGGAAGCCCAACACACAGCTTTGACGAGGAAACAGAAAAGGGAG 1440
QY 1441 AGATTTCTGGGCATTAAGGAGGCCACAGAAAGAGAGCCAGGGCCCCCAAGTCTCTCTTT 1500
Db 1441 AGATTTCTGGGCATTAAGGAGGCCACAGAAAGAGAGCCAGGGCCCCCAAGTCTCTCTTT 1500
QY 1501 ATACCTCTCATCCCGTCTCCCAATTAAAGCCACTCTTTCTCTTAGATCAGACCTGAGCTGC 1560
Db 1501 ATACCTCTCATCCCGTCTCCCAATTAAAGCCACTCTTTCTCTTAGATCAGACCTGAGCTGC 1560
QY 1561 AGCGAAGAGACCCCTGAGGAGGATCACTTGGATGAAGAGATGTGTGAGAAAGTCCAGG 1620
Db 1561 AGCGAAGAGACCCCTGAGGAGGATCACTTGGATGAAGAGATGTGTGAGAAAGTCCAGG 1620
QY 1621 GCAACCTTAAGAGCCAGGCTTAAAGCAAGAGATAAAGGTGCTTCAAGGTGGCCAGG 1680
Db 1621 GCAACCTTAAGAGCCAGGCTTAAAGCAAGAGATAAAGGTGCTTCAAGGTGGCCAGG 1680
QY 1681 CTGTGCACACAGAGGCTGAGGACTTGGTGTAGAGCTTCAAGATGAAGGATGATGCTCAGA 1740
Db 1681 CTGTGCACACAGAGGCTGAGGACTTGGTGTAGAGCTTCAAGATGAAGGATGATGCTCAGA 1740
QY 1741 ATGGCGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGTGAGAAAGAGCCCTGGAACAG 1800
Db 1741 ATGGCGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGTGAGAAAGAGCCCTGGAACAG 1800
QY 1801 AGAATCTGGAAGCGCTGGAACACGATACATAAAGGGAAGAACCCAGGCTTACCTTTAGATG 1860
Db 1801 AGAATCTGGAAGCGCTGGAACACGATACATAAAGGGAAGAACCCAGGCTTACCTTTAGATG 1860
QY 1861 TAAATCATCAAGACAGGGAGGAGGAGCTGGAGAGTAGAAGACCCCGGGGCAAGA 1920
Db 1861 TAAATCATCAAGACAGGGAGGAGGAGCTGGAGAGTAGAAGACCCCGGGGCAAGA 1920
QY 1921 CATGGAAGCAGGACCAAGCCAGCTTTCAGCGCTCCGTGAAATCAGCGCTGTGAAGGACAGAG 1980
Db 1921 CATGGAAGCAGGACCAAGCCAGCTTTCAGCGCTCCGTGAAATCAGCGCTGTGAAGGACAGAG 1980
QY 1981 CCTCGGTATGAGCAACAGAACAGCAGAGGCTAGGGTTAATGTGAGACAGAGGAAACAGAAAG 2040
Db 1981 CCTCGGTATGAGCAACAGAACAGCAGAGGCTAGGGTTAATGTGAGACAGAGGAAACAGAAAG 2040
QY 2041 CTGAGACAGAGGACAGACAGCGGGGAGCCAGGTAAAGGAATGCTCTCTCTCAC 2100
Db 2041 CTGAGACAGAGGACAGACAGCGGGGAGCCAGGTAAAGGAATGCTCTCTCTCAC 2100
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCACATATCTTAGAATGTTTATCAGACTGCAAGGCG 2160
Db 2101 CTGTGGCCAGAGCGTCCATCTGTGTCACATATCTTAGAATGTTTATCAGACTGCAAGGCG 2160
QY 2161 TGGCTTGGAGGACAGCTGGAAAGATGTGTGAGCCAGGGGAGACAAGGGGGCTTAGGA 2220
Db 2161 TGGCTTGGAGGACAGCTGGAAAGATGTGTGAGCCAGGGGAGACAAGGGGGCTTAGGA 2220
QY 2221 AAGGAAGAGAGGGCAAAACCCAGGCCACACAGAGGCGAGGCCAGAACTGAGTTAACTC 2280
Db 2221 AAGGAAGAGAGGGCAAAACCCAGGCCACACAGAGGCGAGGCCAGAACTGAGTTAACTC 2280
QY 2281 CTTCCTCTGTGCTATCTCCATAGGAGGAGTGGGAACCTGTGACCCACCATCCCCATGA 2340
Db 2281 CTTCCTCTGTGCTATCTCCATAGGAGGAGTGGGAACCTGTGACCCACCATCCCCATGA 2340
QY 2341 GCGCCCACTTACCATAACCAAGTTTGGCCCTGAGTGGCAITTTAGGTTTCCCTGAGGACAGAG 2400
Db 2341 GCGCCCACTTACCATAACCAAGTTTGGCCCTGAGTGGCAITTTAGGTTTCCCTGAGGACAGAG 2400
QY 2401 CTTGCGCTTTGTCTCTTGAACCTGACCCAGCTGACCCCAATGTTTCTCAGTACCTTTGAAT 2460
Db 2401 CTTGCGCTTTGTCTCTTGAACCTGACCCCAATGTTTCTCAGTACCTTTATCAT 2460
QY 2461 GCGCTCAAGAGTTCAGAACAGGAGGAGTACATATTTAGCCATGGCTTAAACCTGGAGCT 2520

Db 2461 GCCCTCAAGAGCTTGAGAACACGAGCAGTGCACATATTAGGCCATGGGCTAAACCCCTGGAGCT 2520
Qy 2521 TGCACACAGGAGCCTCAAGTGACCTCCAGGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
Db 2521 TGCACACAGGAGCCTCAAGTGACCTCCAGGGACACAGCTGCAGACAGGTGGCTTTATCC 2580
Qy 2581 CCNAAGAGCAACCATTTGGCTAGGTGGCTGCAAAATGGGAAATGCAAGGTGGAATCAGGTG 2640
Db 2581 CCNAAGAGCAACCATTTGGCTAGGTGGCTGCAAAATGGGAAATGCAAGGTGGAATCAGGTG 2640
Qy 2641 CCTTCAAGAACTACTGCATGCAAGACCTAAGACCCCTGGAGAGAGGGGTATGCTCTGCCC 2700
Db 2641 CCTTCAAGAACTACTGCATGCAAGACCTAAGACCCCTGGAGAGAGGGGTATGCTCTGCCC 2700
Qy 2701 CCACCCACCAAGGGGAGTGAACTATCCTAGGGGCTGGGAGCTTGGGGAGACACAC 2760
Db 2701 CCACCCACCAAGGGGAGTGAACTATCCTAGGGGCTGGGAGCTTGGGGAGACACAC 2760
Qy 2761 ATTACTGAGAGTGTGAGGCCAGAAAACTGACCCGCCCTGTGCTGCTGCCACCTCCACAC 2820
Db 2761 ATTACTGAGAGTGTGAGGCCAGAAAACTGACCCGCCCTGTGCTGCTGCCACCTCCACAC 2820
Qy 2821 TCTAGAGCTATATTGAGAGTGCATGATAGATAGGTGGAGCTGTAGCAGGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTGAGAGTGCATGATAGATAGGTGGAGCTGTGTAGCAGGAGAGTGT 2880
Qy 2881 TCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTGTCTCTGAA 2940
Db 2881 TCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTGTCTCTGAA 2940
Qy 2941 CACAATGTCTACTTGTATTAAAGGCACTGACTCTAAAGACCCAAACATCTACGACCTC 3000
Db 2941 CACAATGTCTACTTGTATTAAAGGCACTGACTCTAAAGACCCAAACATCTACGACCTC 3000
Qy 3001 TGAAGAAGACAGAGCCCTGGAGACAGGGTGTCTCTGAGCCTTGGGTCTGTATGTTG 3060
Db 3001 TGAAGAAGACAGAGCCCTGGAGACAGGGTGTCTCTGAGCCTTGGGTCTGTATGTTG 3060
Qy 3061 CCACAAAGGAGGCATGATGTGAGTATAGGCCCCAGGAGGGTTAGAGAGGGCACTTG 3120
Db 3061 CCACAAAGGAGGCATGATGTGAGTATAGGCCCCAGGAGGGTTAGAGAGGGCACTTG 3120
Qy 3121 GGAAGGGTGTGCTGCGAGAGCCCTATCCATGGAATCTGAGCCTGGGGCCCACTGTTG 3180
Db 3121 GGAAGGGTGTGCTGCGAGAGCCCTATCCATGGAATCTGAGAGCCTGGGGCCCACTGTTG 3180
Qy 3181 TAAATCTCTGGCCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCTGGCAGCCTGGGGA 3240
Db 3181 TAAATCTCTGGCCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCTGGCAGCCTGGGGA 3240
Qy 3241 GSCGGAAGGAGCAACCCGCCACTTATACCCCTTCTCCCTCAGCCCCCAGGATTACACCT 3300
Db 3241 GSCGGAAGGAGCAACCCGCCACTTATACCCCTTCTCCCTCAGCCCCCAGGATTACACCT 3300
Qy 3301 CTGGCCTTCCCTTCCCACTCCCATCAGGAGTGGAGGGTTGCAGAGGGAGGGTAAAAA 3360
Db 3301 CTGGCCTTCCCTTCCCACTCCCATCAGGAGTGGAGGGTTGCAGAGGGAGGGTAAAAA 3360
Qy 3361 CCTACATGTCCAAAATCATGTGTGACGATATATGATATATGATATGATGATGATGATGATG 3420
Db 3361 CCTACATGTCCAAAATCATGTGTGACGATATATGATATATGATATGATGATGATGATGATG 3420
Qy 3421 GGAATCTGCAGGCTTAACCTGGGTTAATGTGTAAGTCTGTGTGATGATGATGATGATGATG 3480
Db 3421 GGAATCTGCAGGCTTAACCTGGGTTAATGTGTAAGTCTGTGTGATGATGATGATGATGATG 3480
Qy 3481 ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAAGTCTGTGCGTGAAGTTACCAGACTGCA 3540
Db 3481 ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAAGTCTGTGCGTGAAGTTACCAGACTGCA 3540
Qy 3541 GGTGTTGTGTAATTTGCCAAAGGCAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
Db 3541 GGTGTTGTGTAATTTGCCAAAGGCAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600

Qy 3601 TGGATGATGGCTGCATCTCAAGGACCATGAAAAATAGAAATGGACACTCTATATGTGTCT 3660
Db 3601 TGGATGATGGCTGCATCTCAAGGACCATGAAAAATAGAAATGGACACTCTATATGTGTCT 3660
Qy 3661 CTAAGCTAAGGTAGCAAGGTCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
Db 3661 CTAAGCTAAGGTAGCAAGGTCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
Qy 3721 TACAGACATGATCTGTACAGAGTAAAGGAGAGAGAGGGGTGTAGAAATTTCTTACTTA 3780
Db 3721 TACAGACATGATCTGTACAGAGTAAAGGAGAGAGAGGGGTGTAGAAATTTCTTACTTA 3780
Qy 3781 TCAAAAGGAAACTCAGTCTGCACCTGCAAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840
Db 3781 TCAAAAGGAAACTCAGTCTGCACCTGCAAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840
Qy 3841 TGTCTCTGGGAGCCAGCCTGTGGAACCTTCAAGGTCTGAGAGATGAGAGGCTCCCTCA 3900
Db 3841 TGTCTCTGGGAGCCAGCCTGTGGAACCTTCAAGGTCTGAGAGATGAGAGGCTCCCTCA 3900
Qy 3901 GCCTGAAGCTATGAGATAGCCAGGTTTGAAGGGGGAAGGAGAGCCTGGGATGGGAGC 3960
Db 3901 GCCTGAAGCTATGAGATAGCCAGGTTTGAAGGGGGAAGGAGAGCCTGGGATGGGAGC 3960
Qy 3961 TTGTGTGTTGGAGCAGGGACAGATATTAAGCCTTGAAGAGAGAGTGCACCTTACCCAG 4020
Db 3961 TTGTGTGTTGGAGCAGGGACAGATATTAAGCCTTGAAGAGAGAGTGCACCTTACCCAG 4020
Qy 4021 TTGTTCAACTCACCTTTCAGATTAAAAATACTGAGGTAAAGGCTGGGTAGGGAGGTG 4080
Db 4021 TTGTTCAACTCACCTTTCAGATTAAAAATACTGAGGTAAAGGCTGGGTAGGGAGGTG 4080
Qy 4081 GTGTGAGACGCT 4140
Db 4081 GTGTGAGACGCT 4140
Qy 4141 GCCCAGAGCTTAAAAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 4200
Db 4141 GCCCAGAGCTTAAAAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 4200
Qy 4201 GCAAACTTGGGGCCCTGTGATCGATTGCAAGAACTCGCCCAATCGATACCTCTCTTC 4260
Db 4201 GCAAACTTGGGGCCCTGTGATCGATTGCAAGAACTCGCCCAATCGATACCTCTCTTC 4260
Qy 4261 TTCTAAAGGACAGAGGGAACTCGAGTTTACATCTCTCTCTCTCTCTCTCTCTCTCTCT 4320
Db 4261 GGAGCCAGGACAGAGGGAACTGGGGAGGAG----- 4291
Qy 4321 AAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTTCGAGTTTACCCTC 4380
Db 4292 ----- 4291
Qy 4381 CCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAA 4440
Db 4292 ----- 4291
Qy 4441 AAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAGTTCGAGTTT 4500
Db 4292 ----- 4291
Qy 4501 CCACTCCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGAT 4560
Db 4292 ----- 4291
Qy 4561 AGAGAAAAGTGAAGTTCGAGTCTCGGTCCAGCAGAGGACTCCAAATTTAGGAGGAGGCA 4620
Db 4292 -----GGTCCAGCAGAGGACTCCAAATTTAGGAGGAGGCA 4328
Qy 4621 TATGGATGGATATAAAGGGGCTGGAGCACTGAGAGCTGTGAGAGATTTCTTCAACCCA 4680
Db 4329 TATGGATGGATATAAAGGGGCTGGAGCACTGAGAGCTGTGAGAGATTTCTTCAACCCA 4388

QY 4681 GGTAAGAGGAGCTTTTCGGGTGGGGCTCTTTACCCACACAGACCTCTCCACCTAGAA 4740
DB 4389 GGTAAGAGGAGCTTTTCGGGTGGGGCTCTTTACCCACACAGACCTCTCCACCTAGAA 4448
QY 4741 GGAATCTGCTTTCTCTGGAAGTGGGTTTCAGCCGGTTCAGAGATCTGACAGGTTGCCCTT 4800
DB 4449 GGAATCTGCTTTCTCTGGAAGTGGGTTTCAGCCGGTTCAGAGATCTGACAGGTTGCCCTT 4508
QY 4801 CCACGAGCTGGAGATCTCAGTGGCAGGAGTTTCCACAGAAACACTGATGCCCTT 4860
DB 4509 CCACGAGCTGGAGATCTCAGTGGCAGGAGTTTCCACAGAAACACTGATGCCCTT 4568
QY 4861 TCCCTTAGGCTCTCTCTCCATCTCTCTCTGGGATGCTCTCTCCCGTCTGGTTATC 4920
DB 4569 TCCCTTAGGCTCTCTCTCCATCTCTCTCTGGGATGCTCTCTCCCGTCTGGTTATC 4628
QY 4921 TTGGCTCTCTCTCTCAGCAAGATTTGCCCTGTGTGTCATCTCCATCTTTCTCTACTGT 4980
DB 4629 TTGGCTCTCTCTCTCAGCAAGATTTGCCCTGTGTGTCATCTCCATCTTTCTCTACTGT 4688
QY 4981 CTCCTGTCTTGCCTTGCCTTCTTGGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5040
DB 4689 CTCCTGTCTTGCCTTGCCTTCTTGGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4748
QY 5041 TTTCTCTCCCTTCTCATTTGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5100
DB 4749 TTTCTCTCCCTTCTCATTTGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4808
QY 5101 CT 5160
DB 4809 CT 4868
QY 5161 CTGTGTCTGAGTCTGAGATCACACCTGGGTTCACCCCTTATGTAAACATCTTCCA 5220
DB 4869 CTGTGTCTGAGTCTGAGATCACACCTGGGTTCACCCCTTATGTAAACATCTTCCA 4928
QY 5221 GTGAGCCACAGTCTCAGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5280
DB 4929 GTGAGCCACAGTCTCAGTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4988
QY 5281 CTGTCTCCATCTCTGCTCAGGATCTCTAGATTTGCTCTCCAGCTCTCTCTCTCTCTCT 5340
DB 4989 CTGTCTCCATCTCTGCTCAGGATCTCTAGATTTGCTCTCCAGCTCTCTCTCTCTCTCT 5048
QY 5341 GCTGTCT 5400
DB 5049 GCTGTCT 5108
QY 5401 ATTCTTCAGGATCTCTGAAAGTTTAAACAGGTGAGATTTTCCCTGTAGACAGCAGA 5460
DB 5109 ATTCTTCAGGATCTCTGAAAGTTTAAACAGGTGAGATTTTCCCTGTAGACAGCAGA 5168
QY 5461 TCACGATCTCTCCGGAAGTCCAGCTTCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5520
DB 5169 TCACGATCTCTCCGGAAGTCCAGCTTCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5228
QY 5521 CTCTTAGCAAACTCAGGACCTTACCCCAATAGACCTCTGACAGAGAAGCAGCACT 5580
DB 5229 CTCTTAGCAAACTCAGGACCTTACCCCAATAGACCTCTGACAGAGAAGCAGCACT 5288
QY 5581 TTACATGGATCTCTGTTGGAGGATGCTAGGCTAGGCTGTTAAAGGAGGAGGAGTGTGT 5640
DB 5289 TTACATGGATCTCTGTTGGAGGATGCTAGGCTAGGCTGTTAAAGGAGGAGGAGTGTGT 5348
QY 5641 GGTGTAGGAAGTCCAGACTCTACATAGAGCTTACCCACACAGCAATGACAGACAGA 5700
DB 5349 GGTGTAGGAAGTCCAGACTCTACATAGAGCTTACCCACACAGCAATGACAGACAGA 5408
QY 5701 TCCCTCTCTATCTCCCCCATAAGATTTGAGTCGAC 5735
DB 5409 TCCCTCTCTATCTCCCCCATAAGATTTGAGTCGAC 5443

RESULT 3
AC09577
LOCUS
DEFINITION
AC09577
VERSION
KEYWORDS
SOURCE
ORGANISM

AC09577 235759 bp DNA linear HTG 22-FEB-2004
Mus musculus chromosome 14 clone RP23-171A13 map 14, *** SEQUENCING
IN PROGRESS ***, 6 unordered pieces.
AC09577
HTG: HTGS, PHASE1; HTGS, FULLTOP; HTGS, ACTIVEFIN.
AC09577
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 235759)
Birren, B., Nussbaum, C. and Lander, E.

Mus musculus chromosome 14, clone RP23-171A13

Unpublished

2 (bases 1 to 235759)

Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lechoczky, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 235759)

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Submitted (22-FEB-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 22, 2004 this sequence version replaced gi:31581779.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L16015
 Center clone name: 171_A_13

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 21715: contig of 21715 bp in length
 * 21716 21815: gap of 100 bp
 * 21816 53733: contig of 31918 bp in length
 * 53734 53833: gap of 100 bp
 * 53834 113500: contig of 59667 bp in length
 * 113501 113600: gap of 100 bp
 * 113601 196948: contig of 83348 bp in length
 * 196949 197048: gap of 100 bp
 * 197049 205662: contig of 8614 bp in length
 * 205663 205762: gap of 100 bp
 * 205763 235759: contig of 29997 bp in length.

FEATURES

source

1..235759
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="14"
 /map="14"
 /clone="RP23-171A13"
 /clone_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 83.8%; Score 4804.8; DB 2; Length 235759;
 Best Local Similarity 92.0%; Pred. No. 0;
 Matches 5316; Conservative 0; Mismatches 107; Indels 357; Gaps 9;
 QY 1 GGATCCTGCAAGGTCAACAAGGGTCTCCACCCACAGGTGCCCTAGTCTCAATTTAGT 60
 DB 211665 GGATCCTGCAAGGTCAACAAGGTCTCCACCCACAGGTGCCCTAGTCTCAATTTAGT 211724
 QY 61 TTCATGCTTGTCTCACAATGCTGGCTCCCGAGAGTAATTTGGACTTTGTTTAT 120
 DB 211725 TTCATGCTTGTCTCACAATGCTGGCTCCCGAGAGTAATTTGGACTTTGTTTAT 211784
 QY 121 TTCAAAAGGGCTGAATGAGGATAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCGTGA 180
 DB 211785 TTCAAAAGGGCTGAATGAGGATAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCGTGA 211844
 QY 181 AGCCCTCAGACTGAGGCTTTGCAACAGCCCTTTAGTGGAGCAGATAAAGCAATTT 240
 DB 211845 AGCCCTCAGACTGAGGCTTTGCAACAGCCCTTTAGTGGAGCAGATAAAGCAATTT 211904
 QY 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTCTGACCTCGGTCCCTGGGCTCT 300
 DB 211905 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTCTGACCTCGGTCCCTGGGCTCT 211964
 QY 301 AGGGTGGGAGGTGGGGCTTGAAGAAGAGTGGGAAGTGGCAAAAGCCGATCCCTAG 360
 DB 211965 AGGGTGGGAGGTGGGGCTTGAAGAAGAGTGGGAAGTGGCAAAAGCCGATCCCTAG 212024
 QY 361 GGCCCTGTGAATTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCACGCC 420
 DB 212025 GGCCCTGTGAATTCGGAGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCACGCC 212084
 QY 421 AAACATAGCAAGATGATACCTCTTTGTGACTTTCCCGAGCCAGCTACCTGTGAGTT 480
 DB 212085 AAACATAGCAAGATGATACCTCTTTGTGACTTTCCCGAGCCAGCTACCTGTGAGTT 212144
 QY 481 GAAACAGGATTTAGAGAAGCCTCTGAACCTCACCTGAACTCTGAAAGTATCCACCAAGCA 540

DB 212145 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAACTCTGAAAGTATCCACCAAGCA 212204
 QY 541 AGCACTAGTGGCCACTGCTAGTATTAGTATCTTACGCTGATATATGAGAGTGGGCCAC 600
 DB 212205 AGCACTAGTGGCCACTGCTAGTATTAGTATCTTACGCTGATATATGAGAGTGGGCCAC 212264
 QY 601 AGAAGTCTCGGGGTAGGAACTGACAGTCACTTTTTCAGTCGGCAAGGATATGACCCCC 660
 DB 212265 AGAAGTCTCGGGGTAGGAACTGACAGTCACTTTTTCAGTCGGCAAGGATATGACCCCC 212324
 QY 661 TCACAGATGTAGTAATGTCTCCCTTAGATCCCATCCAGGAGGTCTCTAAGAGGACATG 720
 DB 212325 TCACAGATGTAGTAATGTCTCCCTTAGATCCCATCCAGGAGGTCTCTGAGAGGACAG 212384
 QY 721 GGATGAGAGATGTAGTCAATGTGGCATTTCCAAACACAGCATATCCAGTGTCTCTGGCCC 780
 DB 212385 GGATGAGAGATGTAGTCAATGTGGCATTTCCAAACACAGCATATCCAGTGTCTCTGGCCC 212444
 QY 781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCTCATCTTTCTCTCCCTCCCTCCCTCCAG 840
 DB 212445 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCTCATCTTTCTCTCCCTCCCTCCCTCCAG 212504
 QY 841 GACACACCCCTGCTGTCAGTATTCAATTTCTTCCCTTCCAGTCCCTCTGTGACTTCCAT 900
 DB 212505 GACACACCCCTGCTGTCAGTATTCAATTTCTTCCCTTCCAGTCCCTCTGTGACTTCCAT 212564
 QY 901 TTGCAAGGCTTTTGAACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGTGTGGCAA 960
 DB 212565 TTGCAAGGCTTTTGAACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGTGTGGCAA 212624
 QY 961 GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGATTTTGGAAAGATCAGAACTAAA 1020
 DB 212625 GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGATTTTGGAAAGATCAGAACTAAA 212684
 QY 1021 TCACCTGGCGGCTTGGGGGTAGAAAAAGAGTGAAGTGAAGTCCGCTTCCAGCTTAAGCAAGC 1080
 DB 212685 TCACCTGGCGGCTTGGGGGTAGAAAAAGAGTGAAGTGAAGTCCGCTTCCAGCTTAAGCAAGC 212744
 QY 1081 TAGTCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTCT 1140
 DB 212745 TAGTCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTCT 212804
 QY 1141 TGAAGCAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCACCTACAGAC 1200
 DB 212805 TGAAGCAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCACCTACAGAC 212864
 QY 1201 ACACCTGCTGTGGCCAGACTCTCTGTTCAACAGCCCTCTGTGTTCTGACCCTAGAGTAG 1260
 DB 212865 ACACCTGCTGTGGCCAGACTCTCTGTTCAACAGCCCTCTGTGTTCTGACCCTAGAGTAG 212924
 QY 1261 GCAACAGAGATGGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
 DB 212925 GCAACAGAGATGGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 212984
 QY 1321 CAGGGGGGAGAACAGAGAACAAATAAGGAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
 DB 212985 CAGGGGGGAGAACAGAGAACAAATAAGGAAGAAAGGAGGAGGAGGAGGAGGAGGAG 213044
 QY 1381 TGCAGTCAGAAAGATGGGAAGCCAAACACACAGCTTTGAGCAGAGGAAACAGAAAAAGGAG 1440
 DB 213045 TGCAGTCAGAAAGATGGGAAGCCAAACACACAGCTTTGAGCAGAGGAAACAGAAAAAGGAG 213104
 QY 1441 AGATTCTGGGATTAAGGAGGCCACAGAAAGAGGCCAGGCCGCCGCCAAAGTCTCTCTTT 1500
 DB 213105 AGATTCTGGGATTAAGGAGGCCAAAGAAAGAGGCCAGGCCGCCGCCAAAGTCTCTCTTT 213164
 QY 1501 ATACCTCATCCCTCTCCCAATTAAGCCCACTCTTCTTCTAGATCAGACTGAGTGC 1560
 DB 213165 ATACCTCATCCCTCTTCCAAATTAAGCCCACTCTTCTTCTAGATCAGACTGAGTGC 213224
 QY 1561 AGCAGAGAGCCGTTAGGAGGATCACACTGGATGAAGAGATGTGTGGAGAGTCCAGG 1620

| | | | | | | | |
|----|--------|--|--------|----|--------|---|--------|
| Db | 213225 | AGCGAAGAGACCCGTAGGAGGATCACACTGGATGAAGGAGATGTGTGAGAAAGTCCAGG | 213284 | QY | 2697 | GCCCCACCCACCCATAAGGGGAGTGAACCTATCTAGGGGGCTGGCGACCTTGGGGAGACA | 2756 |
| QY | 1621 | GCACCTTAGAGCCAGAGCCTAAAGAGCAAGATAAAGTGTCTTCAAGGTGGCCAGG | 1680 | Db | 214352 | GCCCCACCCACCCATAAGGGGAGTGAACCTATCTAGGGGGCTGGCGACCTTGGGGAGACA | 214411 |
| Db | 213285 | G-AACCTAAGAGCCAGAGCCTAAAGAGCAAGATAAAGTGTCTTCAAGGTGGCCAGG | 213343 | QY | 2757 | CCACATTACTGAGAGTGTGAGCCCGAGAAAACTGACCGCCCTGTGTCTCTGCCACCTCC | 2816 |
| QY | 1681 | CTGTGCACACAGAGGCTGAGAGCTGGTGGTAGAGCCTCAAGATGAAGATGATGTCTCAGA | 1740 | Db | 214412 | CCATATTACTGAGAGTGTGAGCCCGAGAAAACTGACCGCCCTGTGTCTCTGCCACCTCC | 214471 |
| Db | 213344 | CTGTGCACACAGAGGCTGAGAGCTGGTGGTAGAGCCTCAAGATGAAGATGATGTCTCAGA | 213403 | QY | 2817 | ACACTCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGGAGCTGTAGCAGGGAGA | 2876 |
| QY | 1741 | ATGGGGGGGGGG- GGGGATTTCTGGGGGGGGGAGAGAAAGGTGAAGAGAGCTGTGAACA | 1799 | Db | 214472 | ACACTCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGGAGCTGTAGCAGGGAGA | 214531 |
| Db | 213404 | ATGGGGGGGGGGTGGGGGGAGTCTGGGGGGGGGAGAGAAAGGTGACAGAGAGCTGTGAACA | 213463 | QY | 2877 | GTGTTCTCTGGGTGTGAGGGGTAGGGGAAAGCAGAGCAGGGAGTCTGCGCTTTGTCTCC | 2936 |
| QY | 1800 | GAGATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGACCCAGGCTACTCTTTAGAT | 1859 | Db | 214532 | GTGTTCTCTGGGTGTGAGGGGTAGGGGAAAGCAGAGCAGGGAGTCTGCGCTTTGTCTCC | 214591 |
| Db | 213464 | GAGATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGACCCAGGCTACTCTTTAGAT | 213523 | QY | 2937 | TGAACACAATGTCTACTTAGTTTATAACAGGCATGACCTGTCTAAAGACCCAAACATCTACGA | 2996 |
| QY | 1860 | GTAAATCATGAAGACAGGGAGAAAGGAAAGCTGGAGAGT---AGAAAGACCCCGGGC | 1916 | Db | 214592 | TGAACACAATGTCTACTTAGTTTATAACAGGCATGACCTGTCTAAAGACCCAAACATCTACGA | 214651 |
| Db | 213524 | GTAAATCATGAAGACAGGGAGAAAGGAAAGCTGGAGAGT---AGAAAGACCCCGGGC | 213583 | QY | 2997 | CCTCTGAAAAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTTGGGTGCTTGAT | 3056 |
| QY | 1917 | AAGACATGGAAGCAAGGACCAAGCAGGTTGAGCGCTCCGTGAAATCAGCCTGTGTAAGGC | 1976 | Db | 214652 | CCTCTGAAAAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTTGGGTGCTTGAT | 214711 |
| Db | 213584 | AAGACATGGAAGCAAGGACCAAGCAGGTTGAGCGCTCCGTGAAATCAGCCTGTGTAAGGC | 213643 | QY | 3057 | GGTGCCACANAAGGAGGCGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCA | 3116 |
| QY | 1977 | AGAGCCCTGTTATGAGCACCAGAACAGCAGAGGCTTAGGTTTAATGTTCGAGACAGGAAACA | 2036 | Db | 214712 | GGTGCCACANAAGGAGGCGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCA | 214771 |
| Db | 213644 | A-----GAGCACAGAACACCAGAGGCTAGGGTTAATGTTCGAGACAGGAAACA | 213691 | QY | 3117 | CTTTGGGAAGGGGTCAGTCTGACAGGCCCTATCCATGGAATCTGAGAGCCTGGGGCCAACT | 3176 |
| QY | 2037 | GAAAGGTAGACACAGGAACAGACAGACGGGGGAGCCAGTAAACAAAGGAATGTCTCTTC | 2096 | Db | 214772 | CTTTGGGAAGGGGTCAGTCTGACAGGCCCTATCCATGGAATCTGAGAGCCTGGGGCCAACT | 214831 |
| Db | 213692 | GAAAGGTAGACACAGGAACAGACAGACGGGGGAGCCAGTAAACAAAGGAATGTCTCTTC | 213751 | QY | 3177 | GGTGTAAATCTCTGGCCTGCGCAGGCAATCAAAGCAGCACCTGCAATCTCTGGAGCCTG | 3236 |
| QY | 2097 | TCACTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCATCAGACTGA | 2156 | Db | 214832 | GGTGTAAATCTCTGGCCTGCGCAGGCAATCAAAGCAGCACCTGCAATCTCTGGAGCCTG | 214891 |
| Db | 213752 | TCACTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCATCAGACTGA | 213811 | QY | 3237 | GGGAGGCGGAAGGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAC | 3296 |
| QY | 2157 | GGGCTGGCTTGGGAGCAGCTGGAAGAGTATGTGAGAGCCAGGGAGACAAGGGGGCTT | 2216 | Db | 214892 | GGGAGGCGGAAGGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAC | 214951 |
| Db | 213812 | GGGCTGGCTTGGGAGCAGCTGGAAGAGTATGTGAGAGCCAGGGAGACAAGGGGGCTT | 213871 | QY | 3297 | ACCTCTGGCCTTCCGCCCTTCCACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGGGTA | 3356 |
| QY | 2217 | AGGAAGGAAGAGGGGCAACCCAGGCCACACAGAGGGCAGAGCCAGAACTCAGTTA | 2276 | Db | 214952 | ACCTCTGGCCTTCCGCCCTTCCACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGGGTA | 215011 |
| Db | 213872 | AGGAAGGAAGAGGGGCAACCCAGGCCACACAGAGGGCAGAGCCAGAACTCAGTTA | 213931 | QY | 3357 | AAAACTACATGTCAAACATCATGGTGACATATATGATCAGTATGTAGAGGCAA | 3416 |
| QY | 2277 | ACTCTTCTCTGTGATCTTCCATAGAGGAGTGGGAACTCTGTGACCAACATCCCCC | 2336 | Db | 215012 | AAAACTACATGTCAAACATCATGGTGACATATATGATCAGTATGTAGAGGCAA | 215071 |
| Db | 213932 | ACTCTTCTCTGTGATCTTCCATAGAGGAGTGGGAACTCTGTGACCAACATCCCCC | 213991 | QY | 3417 | GAAAGGAATCTGACGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGATGTGTGTG | 3476 |
| QY | 2337 | ATGAGCCCCACTACCCATACCAAGTTTGGCTGAGTGGCATCTTAGGTTTCCCTGAGGAC | 2396 | Db | 215072 | GAAAGGAATCTGACGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGATGTGTGTG | 215131 |
| Db | 213992 | ATGAGCCCCACTACCCATACCAAGTTTGGCTGAGTGGCATCTTAGGTTTCCCTGAGGAC | 214051 | QY | 3477 | TCGTGCTGAAAAACGGGCATGGCTGTGACGCTGTTCAGTTCTGTGCTGAGGTTTACAGAC | 3536 |
| QY | 2397 | AGAGCTGGCTTGTCTCTTGACCTGACCCAGCTGACCCAAATGTTCTCAGTACCTTT | 2456 | Db | 215132 | TCGTGCTGAAAAACGGGCATGGCTGTGACGCTGTTCAGTTCTGTGCTGAGGTTTACAGAC | 215191 |
| Db | 214052 | AGAGCTGGCTTGTCTCTTGACCTGACCCAAAGCTGACCCAAATGTTCTCAGTACCTTTA | 214111 | QY | 3537 | TGCAGGTTTGTGTGTAATTTGCCAAGGCAAGTGGTGAATCCCTTCCATGGTTTAAAG | 3596 |
| QY | 2457 | GAAATGCCCTCAAGAGCTTGAGAACACAGGAGTGAATATAGGCCATGGGCTTAACCCCTGG | 2516 | Db | 215192 | TGCAGGTTTGTGTGTAATTTGCCAAGGCAAGTGGTGAATCCCTTCCATGGTTTAAAG | 215251 |
| Db | 214112 | TCATGCCCTCAAGAGCTTGAGAACACAGGAGTGAATATAGGCCATGGGCTTAACCCCTGG | 214171 | QY | 3597 | AGATTGGATGATGGCCTGCATCTCAAGGACCAATGGAATAAGATGGAACCTCTATATGT | 3656 |
| QY | 2517 | AGCTTGACACAGGAGCTCAAGTGAACCTCCAGGACACAGCTGACAGAGGTGGCTTTT | 2576 | Db | 215252 | AGATTGGATGATGGCCTGCATCTCAAGGACCAATGGAATAAGATGGAACCTCTATATGT | 215311 |
| Db | 214172 | AGCTTGACACAGGAGCTCAAGTGAACCTCCAGGACACAGCTGACAGAGGTGGCTTTT | 214231 | QY | 3657 | GTCTCTAAAGTAAAGTACAGGCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAG | 3716 |
| QY | 2577 | ATCCCCAAGAGCAACATTTGGCATAGTGGCTGCAATGGAATGGAATGCAAGTTGAATCA | 2636 | Db | 215312 | GTCTCTAAAGTAAAGTACAGGCTTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAG | 215371 |
| Db | 214232 | ATCCCCAAGAGCAACATTTGGCATAGTGGCTGCAATGGAATGGAATGCAAGTTGAATCA | 214291 | QY | 3717 | AGACTACAGACATGATCTGTACAGAGTAAAGAGAGAGAGAGGGG-GTGTAGAAATTTCTCT | 3775 |
| QY | 2637 | GTTCCTTCAAGAACTGTCATGACAGACCTTAAGACCCCTGGAGAGAGGGGTATGTCCT | 2696 | Db | 215372 | AGACTACAGACATGATCTGTACAGAGTAAAGAGAGAGAGAGGGGAGGGGTATGTCCT | 215431 |
| Db | 214292 | GTTCCTTGGAGAACTGTCATGCAAGACCTTAAGACCCCTGGAGAGAGGGGTATGTCCT | 214351 | | | | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

1 (bases 1 to 185702)
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buahy, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huiyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louised, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoileme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzos, W., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

TITLE JOURNAL

REFERENCE
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS

3 (bases 1 to 185702)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23195973.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KCAL
Center clone name: CH230-249H16
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 158880 bases at least Q40
Consensus quality: 159964 bases at least Q30
Consensus quality: 160601 bases at least Q20
Estimated insert size: 162428; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 185702: contig of 185702 bp in length.

FEATURES source

1..185702
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-249H16"
misc_feature
2778..83571
/note="clone_boundary
clone_end:17
site:
end_sequence:BZ214368"
misc_feature
180811..182798
/note="wgs_contig"
misc_feature
184218..185702
/note="wgs_contig"

ORIGIN

Query Match 40.9%; Score 2343; DB 2; Length 185702;
Best Local Similarity 71.2%; Pred. No. 0;
Matches 4321; Conservative 0; Mismatches 795; Indels 951; Gaps 55;
QY 1 GGATCCTGCAAGGTGACACAAAGGGTCTCCACCACAGGTGCCCTAGTCTCAATTTCACT 60
Db 97021 GGATCCTGCAAGGTGACACAAAGGGTCTCCACCACAGGTGCCCTAGTCTCAATTTCACT 97080
QY 61 TTCATGCTGTTCTTCAATGCTGGGCTCCCGAGAGTAATTTGGACTTTGTTTAT 120
Db 97081 TTCATGCTGTTCTTCAATGCTGGGCTCCCGAGAGTAATTTGGACTTTGTTTAT 97139
QY 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGCTACCGACTTAAGGGTCCCGTCA 180
Db 97140 TTCAAAAGGGCTGAATGAGGAGTAGATCTTGCTACCGACTTAAGGGTCCCGTCA 97199
QY 181 AGCCCTCAGACTGGAGCTTTTGCACAGCCCTTTAGGTGGAGAGCAGATAAGCAATTT 240
Db 97200 AGCCCTCAGACTGGAGCTTTTGCACAGCCCTTTAGGTGGAGAGCAGATAAGCAATTT 97259
QY 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTCTGACCTCGGTCCT-GGGCTC 299

Db 97260 TCCCTCAAGCCAAATCTGTCTCTAGACTCTTCTTCACTGACCTCGTCCCTGGGGCTC 97319
QY 300 TAGGGTGGGAGGTGGGCTTGGGAAGAGAGAGGTGGGAAGTGGCAAGCGGATCCCTTA 359
Db 97320 TAGGGTGGGAGGTGGGATTTG- ---AAGAGGAGGAGGAGTGGCAAGGCCACTCCCA 97376
QY 360 GGGCCCTGTGAAGTTTCGAGGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCTCCAGC 419
Db 97377 GGAATCTGTGAAGTTCTGAGCCTTCCAGTACAGTGTGCTGTGCTCATAGATCTCTCTCCAGC 97436
QY 420 CAAACATAGCAAGAGTATCTCTTTGTGACTTCCCAAGGCCAGTACTGTGAGGT 479
Db 97437 CAAACACAGCAAGAGTGTGCTCTCTTTGTGACTTCCCAAGGCCAGTGTGCTGTGCGGT 97496
QY 480 TGAACACAGATTTAGAGAGGCTCTGAACTGACCTGAACTGGAAGCTGGAAGCTATCCCAAGC 539
Db 97497 TGAACACAGATTTAGAGAGGCTCTGAACTGACCTGGAAGCTCTGAGCCCATCTAC- ---C 97552
QY 540 AAGCACCTAGGTGCCACTGTAGTTAGTATCTACGCTGATAATATGACAGAGCTG- -GGC 597
Db 97553 ACACACCTTACATGCCGTTCTATGATATCTTACATGATATCTACATGATATACGAGAGCTGACCA 97612
QY 598 CACGAAGTCTCTGGGGTGTAGGAACTGACAGTGAATTTTCAGTGGCGCAAGGTATGACC 657
Db 97613 GGGAGAAGGCTTGGAGTGTAGGAACTGGCCAGTGAATTTTCAGTCTGAGAGGTATGACC 97672
QY 658 CCCTCAGCAGAT- ---GTAGTAATGTCCCTTAGATCCCATCCAGGAGGCTCTTAAGAG 714
Db 97673 CCCTTGGCAGATATGCGAGATGGGTCCCTTAGAACCCCATCCAGGAGGCTCTGAGAG 97732
QY 715 GACATGGATGAGAGATGTAGTCAATGTGCGCATTTCCAAACACAGCTATCCACAGTGTCCCT 774
Db 97733 GACACAGATGACAGATGTAGTCAATGCGCATTTCCAAACACAGCTATCCACAGTGAATCT 97792
QY 775 TGCCCTCTTCCAATTAGCAGGAGGACAGTAACCTTAGGCTATCTTTCTCTCCCTCCCATCC 834
Db 97793 TGTCCCTTCCAATTAGCAGGAGGATGCAACCTTAGGCTACCTTTCTTCTCCCATCT 97852
QY 835 TCCCAAGGACACACCCCTCTGCTGAGTATTCATTTCTTCTTCACTGCTCCCTCTGTGAC 894
Db 97853 TCCCAAGGACCAACCCCTCTGCTGAGTATTCATTTCTT- CTTCATTTCCCAACATGAC 97911
QY 895 TTCCATTTGCAAGGCTTTTGACCTCTGAGCTGCTGGAAGATAGAGTTTGGCCCTAGTGT 954
Db 97912 TTCCATTTGCAAGGCTTTTACCTCTGAGTGTGGAAGATGGAATTTGGCTCTATATG 97971
QY 955 TGGCAAGGCTATCTAAGAGAAAGCAGACAAACAGGGGACCAAGTTTGGAAAGATCAGGA 1014
Db 97972 TGGCAAAACATCTAAGAGAAAGCAGGCAACAGGGGACCAAGTTTGTAGG- GGGAC 98030
QY 1015 ACTAAATCATCTGGCGGCTGGGGGTAGAAAAGAGTGTGATGCTCCGCTCCAGCTAAG 1074
Db 98031 ACTAAATCACAGGAGGCTGAGAGGTAGAAAAGGATG- -GCTCCAGCTAAG 98082
QY 1075 CCAAGCTAGTCCCGAGATCTTCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATG 1134
Db 98083 CAAAGCCAGATCTTGGGTATTT- ---CACAGCTGGGCTGCT- GAGGTAGGTTTAGGAATG 98137
QY 1135 TGGGTCTGAAAGCAATGGGATTTGGAAGACATCTTTTGAATCTCCCTCAACCCCACT 1194
Db 98138 TGGCTCTGAAAGCAATGGGATTTGAAACATCTCTTTGAACTTGGCTGCTGTGGGT 98197
QY 1195 ACAGACACATCTGTGTGGCCAGATCTCTGTGTTCAACAGCCCTCTGTGTCTGACCATG 1254
Db 98198 AAATCTCCCTCAACCCCAACCCATAGAC- -ACGCTTGTGATGGCCAGACTG 98247
QY 1255 AGCTAGGCAACACAGAGCTGGGCTGTGCTGAGGATGAGAGTTGTTTACCAATAGCAA 1314
Db 98248 AGTTAGGCAACCTGGAGCTGGGCTGTG- TGAGGACGAGG- ---GGTTGCCAATAGCAA 98302
QY 1315 AAACAGCAGGGGAGGAGAACAGAAACGAAATAAGGAGGAGGAAAGGCCAGTCA 1374

Db 98303 AAACAGCAGGTGAGGGAGAAACAGAAAAATGAAATAGGAAGGAGACGGAACGGCC- ---CA 98359
QY 1375 ATCAGATGCACTCAGAGAGATGGGAAGCCAAACACACAGACTTTGAGCAGAGGAAACAGAAA 1434
Db 98360 ATCAGATGCACTCAGAGAGATGGGAACCAATGCATAGCTTTGAAACAAAGAAATAGAG 98419
QY 1435 AGGAGAGAG- - - - - 1443
Db 98420 AGGAAGAGATGGCCAGATCATTAGGTTGGACTTCTCCATTAATTAAGATGCCTCTGTGT 98479
QY 1444 - - - - - TTCTGGGCATAGGAGGCCACAGAAAGAGAGGCCGAGGCC 1484
Db 98480 CCGACTCTACATCCTCGGTACTGTCTATTGCAACTCAGGAATAGCACACAGAGAACT 98539
QY 1485 CCCAAGTCTCTCTTTATACCTCATCCC- - - - - 1513
Db 98540 GTATATGGCCATTTTCAATTTCTCAATCTCTGAAAGCTTGACACACTTGAACAATTTCTAGTT 98599
QY 1514 - - - - - 1513
Db 98600 TACATTTGTGTACTCTATACAGATATGTCTTCATGGACAGTCTGTGTATGCTCCCTGAAT 98659
QY 1514 - - - - - 1513
Db 98660 ACAAGGTCTTTTTTTTATGAAATCAGTGTGCATTTAAATCAAGTCTGCATAGGAGGCCA 98719
QY 1514 - - - - - GTCTCCCAATTAAGCCCACTCTTCTTCTCTAGATCAGAC 1551
Db 98720 CAGGAAGGCCAGGTCCCGAGGTCTCTCTTTATGCTCTCTCCCTCTAAGCCCATTC 98779
QY 1552 CTGAGCTGCAAGGAGAGACCCGTAGGAGGATCACACTGGATGAAGGAGATGTGTGGAG 1611
Db 98780 CTCTTC- CTGGAACAGACTGAGCAGCAGTGGAGAAATCCGTAGAGTCTAAGAGAGATCA 98838
QY 1612 AAGTCCAGGGCAACCTAAGAGCCAGAGCCTTAAAGAGCAAGAGATAAAGGTGCTTCAAAG 1671
Db 98839 CAGTGGAGGAGATGTGTGGACCAAGAGCCTTAAAGAGCAAGAGATAAAGGAGCAGCAAG 98898
QY 1672 GTGGCAGGCTGTG- -CACACAGAGGTCGAGGACTGTGTGTAGAGCCTCAAGATAAGGA 1729
Db 98899 GCAGCTAGCCCGTGCACACACAGAGGTCGAGGACTAGCGGTAGAGCTCTCAAGATAAGGA 98958
QY 1730 TGATGCTCAGAAATGGGCGGGGGGGATTTCTGGGGGGGAGAGAGAGGTGAGAAGGA 1789
Db 98959 TGATGCTCAGAAATGGGTGAGGAGAGAG- - - - - 98987
QY 1790 GCCTGGAAACAGAGAACTTGGAAAGCGCTGGAAAACGATACATATAAGGGAAGAACCCAGGCT 1849
Db 98988 - - - - - GGGACTCAGACTCTGGAAAGCGCTGGAAAACGATACGTGTAAGGGGAAAACCCAGGCT 99043
QY 1850 ACCTTTAGATCTAAATCATGAAAGACAGAGGAGAGGAGGAGCTGGAGAGAGTAGAAGGACC 1909
Db 99044 ACCTTTAGATGGAAGTCAATGAAAGTCAGGAGAGAGGAGGAGCTGGAGAGAGTAGAGG- - - - 99099
QY 1910 CCGGGGCAAGACATGGAAGCAAGGACCAAGCTTGAGCGCTCCGTGAAATCAGCCTGC 1969
Db 99100 - - - - - AATAGACAGGACAGCCAGGTTGAGTGTCTCTGTGAGAT- - - - - 99139
QY 1970 TGAAGGCAAGCCCTGTGTATGAGCACCAGAAACAGCAGAGGTAGGGTTAAATGTGAGACA 2029
Db 99140 - - - - - TCCGATATGAGCGCCGGAACAGCAGAGGCTAGGGCTGATTTAGAGACA 99187
QY 2030 GGGAAACAGAGGTAGACACAGAGAAACAGACAGAGCGGGGAGCCAGGTAAACAAAGGAATG 2089
Db 99188 GGGAACTGAAGTGGGCAACAGAAACAGACAGAGCGGGGATCCAGGTAAACAAAGGAGATG 99247
QY 2090 GTCTTCTCACTGTGGCCAGAGCCTCATCT- GTGTCCACATACCTCTAGAAATGTTTCATC 2148
Db 99248 GTGCTCTCACTGTGGCCAGAGCGCTCATCTGGGGTCCACTTACTCTAGAACGTTTCATC 99307
QY 2149 AGATGCAAGGCTGCTTGGAGGAGGCTGTGAAAGAGATGTGTGAGAGCCAGGGGAGACAA 2208
Db 99308 AGATGCAAGGCTGCTTGGAGGAGGCTGTGAAAGAGATGTGTGAGAG- CGGGGAGACTA 99366

| | | | | | | | |
|----|--------|--|--------|----|--------|--|--------|
| QY | 2209 | GGGGCCCTAGGAAAGGAGGCGCAACAGGCCACAGGCCACACAGAGGGGAGAGCCAGAA | 2268 | QY | 3287 | CAGGATTAAACACCTCTGGCCTTCCCTTCCCACTCCCATCAGGAGTGGAGGTTGCAG | 3346 |
| DB | 99367 | GGGGACCGGAGGAGGAGGAGGCGCAACAGGCCACACAGAGGGGAGAGCCAGAA | 99426 | DB | 100432 | CAGGATTAAACACCTCTGGCCTTCCCTTCCCACTCCCATCAGGAGTGGAGGTTGCAG | 100491 |
| QY | 2269 | CTGAGTTAACTCTCTTCTTGTGCACTTCCATAGAGGAGTGGGAACCTCTGTGACCAAC | 2328 | QY | 3347 | AGGAGGGTAAAACTCATGTCCAAACATCATGGTGACGATATATGATATGATATGT | 3406 |
| DB | 99427 | CTGANTTAACTCTCTTCTTGTGCACTTCCATAGAGGAGTGGGAACCTCTGTGACCAAC | 99483 | DB | 100492 | AGGAGGGTAAAGAACTAGAT-----AATGTGTGATGATATATGATATGATATGT | 100543 |
| QY | 2329 | CATCCCCCATGAGCCGCCACTACCCATACCAATGATGTTGGCTGAGTGGCAATCTAGTTC | 2388 | QY | 3407 | GTAGAGGCAAGAAAGAAATCTGCA-----GGCTTAACTGGGTTAATGTGTAAA | 3457 |
| DB | 99484 | CAT-----CCCCCACTACCTATACCAAGTTTGGCTGAGTAGCACTAGTTC | 99533 | DB | 100544 | GTAGAGGCAAGAAAGAAATCTGGAAGTAAATGTGTAAAAGTCTCTGTGTGTGCAATGTC | 100603 |
| QY | 2389 | CTGAGGACAGAGCTGGCTTGTCTCTTGGACCTTGACCCAAAGCTGACCAATGTTCTCA | 2448 | QY | 3458 | CTGTGTGATGTGTGTGTCTGACTGAAAAAGCGGCATGGCTGTGCACTGTTTCACTTCT | 3517 |
| DB | 99534 | CTGAGGACAAAGCTGTGTTTGTCTGTTGGACCTTGACCCAGCTGACCAATGTTCTCA | 99593 | DB | 100604 | ATGTGAATGCATGTGTGTGTCTGACTGAAAAGTGGGCATGGATGTGCACTGTTTCACTTCT | 100663 |
| QY | 2449 | GTACCTTTGAATGCCCTCAAGAGCTTTGAGAACCGAGCAGTGCATATATTAGGCCATGGCT | 2508 | QY | 3518 | GTGCGTGAAGTTACAGACTGCAGACTGTTGTGTAAAATTGCCCCAGGCAAGTGGGTGAA | 3577 |
| DB | 99594 | GTGCGCTTATCATGCCCTCAAGAGCTTGAGAACCGAGCAGTGCATATATTAGGCCATGGCT | 99653 | DB | 100664 | GTATGTGACATTTACAGACTGCAGCTGCTGTGTGT-----AAGTCAGTGAG | 100708 |
| QY | 2509 | AACCTTGAGCTTGACACAGAGAGCTCAAGTGAAGCTCAGGGACACAGCTGCAGACAGG | 2568 | QY | 3578 | TCCCTTCCATGGTTTAAAGAGATTGGATGATGGCTGCACTCAAGGACCATGGAATA | 3637 |
| DB | 99654 | AACCTTGAGCTTGACACAGAGAGCTCAAGTGAAGCTCAGGGACACAGCTGCAGACAGG | 99713 | DB | 100709 | ACCCTCCCATGGTTTAAAGAAATTTGGATGATAGCCCTGCACTGTCAGGGAACATTTGAAGTA | 100768 |
| QY | 2569 | TGCGCTTTATCCCAAGAGCAACATTTGGCATAGGTGGCTGCAATGGGAATGCAAGG | 2628 | QY | 3638 | GAATGGACACTCTATATGTGTCTCTAAGCTAAGTAGCAAGGTCTTTTGGAGCACACTGT | 3697 |
| DB | 99714 | TGCGCTTTATCTCGGAGAGCAACATTTGGCATAGGTGGCTGCAATGGGAATGCAAGG | 99773 | DB | 100769 | GAATGTAGACTTATATATGTTCCTAAGCAAGTAGCAAGGTCTTTTGGAGCACACTGT | 100828 |
| QY | 2629 | TTGAATCAGGTCCTTTCAAGAACTATGTCATGCAAGACCTTAAGACCCCTGGAGAGAGGGT | 2688 | QY | 3698 | CTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGAGTAAAGAGAGAGAGA | 3757 |
| DB | 99774 | TTGAACAGGTCCTTTCAAGAACTATGTCATGCAAGACCTTAAGACCCCTGGAGAGAGGGT | 99833 | DB | 100829 | CTAGAGATATGGCAACAGAGACAATAGACAGTATCTGTACAGAGTAA--AGAGAGGAGG | 100886 |
| QY | 2689 | ATGCTCTGCCCCCACCACCAATAGGGAGTGAATCTCTAGGGGGTGGCGACCTTG | 2748 | QY | 3758 | GGGGGTGTAGAAATCTCTTATCTAATCAAGGGAAACTGAGTCTGCACCTGCAAGTGGAT | 3817 |
| DB | 99834 | ATGCTCTGCCCCCACCACCAATAGGGAGTGAATCTCTAGGGGGTGGCGACCTTG | 99893 | DB | 100887 | TGAGGTATAGAAATCTCTTACTATCAAGGGGAACTGAGTCTGACCTGCAAAATGAAT | 100946 |
| QY | 2749 | GGGAGACACCAATTACTGAGAGTGTGAGCCCGAGAAACCTGACCGCCCTGTCTCTGC | 2808 | QY | 3818 | GCTCTCCCTAGACATCATGACTTTTCTCTGGGGAGCCAGCACTGTGGAACCTTCAGGTCT | 3877 |
| DB | 99894 | GAGAGACACCAATCATCCGAGAGTGTCTTACTCAGAAACACTGACCGCTGTGTCTCTAC | 99953 | DB | 100947 | GGCTTCCCTGGACATCATGACTTTTCTCTGGGGAGCCAGCACTGTGGAACCTTCAGGTCT | 101006 |
| QY | 2809 | CCACCTCCACACTCTAGAGCTATATT--GAGAGGTGACAGTAGATAGGTGGGAGTGGT | 2866 | QY | 3878 | GAGAGATAGGAGGCTCCCTCAGCCTGAAGCTATGCAAGTATGCAAGTACGAGGTTGAAA--GGG | 3935 |
| DB | 99954 | CCACCTCCACACTCTAGAGCTATATTGAGAGAGGTGACAGTAGAT--GGGGGAGAGTGG | 10012 | DB | 101007 | --GAGAGTAGGAGGCAACCCCTCAGCCTGAAGCTGTGAGATAGTAGGTTGAAAAGAGG | 101064 |
| QY | 2867 | AGCAGGAGAGTGTCTCTGGGTGTGAGGTGTAGGGAAAGCCAGACAGGGAGTCTGG | 2926 | QY | 3936 | GGAAGGAGAGCTCTGGGATGGAGCTGTGTGTGGAGGAGGAGACAGATATTAAGCTCT | 3995 |
| DB | 100013 | AGCAGGAGAGGCTTCTGGGTGTGGGGTGGGA--GGGAAAGTCAAGACAGGGAGTCTGG | 100071 | DB | 101065 | GAAAGGGGAGGCTGGAATGGAGCTGTGTGTGGAGACAGGGGACAAATATTAGGCC | 101124 |
| QY | 2927 | CTTTGTCTCTGAAACAAATGTCTACTTGTAGTTTATTAACAGGCAATGACCTGTAAGACCCA | 2986 | QY | 3996 | GGAAGAGAGGTGACCCCTTACCACCT--TGTTCAACTCACCTTTCAGATTTAAAATAACTG | 4054 |
| DB | 100072 | CTTTGTCTCTGAAACAAATGTCTACTTGTAGTTTATTAACATGCAATGCCGTGTAAGACCCA | 100131 | DB | 101125 | GTAAAGAGAGGTGACCCCTTACCACCTGTGTGTCTCACTCAGCCTTCAGATTTAAAATAACTA | 101184 |
| QY | 2987 | ACATCTACCACTCTGAAAGAGACAGAGCCCTGGAGGACAGGGTTGTCTCTGAGCCTTG | 3046 | QY | 4055 | AGGTAAGGGCC--TGGGTAGGGGAGGTGGTGTGAGAGCTCTCTCTCTCTCTCTGATG | 4111 |
| DB | 100132 | ACACCTCAAACTCTGAAAGAAACAGAGTCTTGGAGGAGCGGGTTGCCGTTGGGCTTG | 100191 | DB | 101185 | AGGTAAGGGCCATGTGGGTAGGGAGTGGTGTGAGACGGTCTCTCTCTCTCTCTATCTG | 101244 |
| QY | 3047 | GCTGTCTGATGTGCCAAAGGAGGGCATGAGTGTAGTATAGGCCCCAGGAGCGTTA | 3106 | QY | 4112 | CCCTCAGGCCCTTTGGGAGGAGGAATGTGCCCAAGGACTAAAAAGGCCATGAGAGCCA | 4171 |
| DB | 100192 | GCTACTTGTGTGCCAAAGGAGGGCATGATGTAGTATAGGCCCCAGGAGCGTTA | 100251 | DB | 101245 | CCCATCGGCCCTTTGGGAGGAGGAATGTGCCCAAGGACTAAAAAGGCCCTTGGAGCCA | 101304 |
| QY | 3107 | GAGAGGGCACTTGGGAAGGGGTCACTCTGCAAGGCCCTTATCCATGGAATCTGAGGCT | 3166 | QY | 4172 | GAGGGCCGAGGGCAACAGACCTTTTCATGGGCAAACTTTGGGCGCCGTAGTGATCGATTGA | 4231 |
| DB | 100252 | GAGAGGGCACTCTGGGAAGGGGTCACTCTGCAAGGCCCTTATCCATGGAATCTGAGACA | 100311 | DB | 101305 | GAGGGCCGAGGGCAACAGACCTTTTCATGGGCAAACTTTGGGCGCCGTAGTGATCGATTGA | 4231 |
| QY | 3167 | GGGGCCAACTGTGTAAATCTCTGGGCTGGCCAGGCAATTAAGAGAGCACTGCACTCTC | 3226 | QY | 4232 | CAAGAACTCGCCAAATCGATACCTTCTTCTTCTTAACGGACAGGAGGAACCTCGAGTTTAC | 4291 |
| DB | 100312 | GGGGCCAACTGTGTAAATCTCTGGGCTGGCCAGGCAATTAAGAGAGCACTGCACTCTC | 100371 | DB | 101348 | -----TGCTGTCTCTCTGTCACCTCCAGAGCCAGGGAATCAAGGAGGAG | 101392 |
| QY | 3227 | TGGAGCCTGGGAGCGGAGAGGAGCAACCCCTTATACCTTTCTCTCTCCTCAGGCC | 3286 | QY | 4292 | CACTCCCTATCAGTGATAGAAAAAGTGAAGTTCGAGTTTACCCTCTCCCTATCAGTGATA | 4351 |
| DB | 100372 | TGGAGCCTGGGAGCGGAGAGGAGCAACCCCTTATACCTTTCTCTCTCCTCAGGCC | 100431 | DB | 101393 | GAGCCAGGACAGGAGGAGTGGAGGGAG-----TGCTGTCTCTCTGTCACCTCCAGAGCCAGGGAATCAAGGAGGAG | 101420 |
| | | | | QY | 4352 | GAGAAAAAGTGAAGTTCGAGTTTACCCTCTCTATCAGTGATAGAGAAAAAGTGAAGTCA | 4411 |

| | | | |
|----|--------|--|--------|
| Db | 101421 | ----- | 101420 |
| Qy | 4412 | GTTTACCACTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCCTCCCTATCA | 4471 |
| Db | 101421 | ----- | 101420 |
| Qy | 4472 | GTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAA | 4531 |
| Db | 101421 | ----- | 101420 |
| Qy | 4532 | AGTCGAGTTTACCCTCCCTATCAGTATAGAGAAAAGTGAAAGTCGAGTCGATACGAG | 4591 |
| Db | 101421 | ----- | 101428 |
| Qy | 4592 | CAGAGGACTCCAAATTTAGGACGACGAGGCATATGGATGGGATATAAAGGGGCTGAGCAC | 4651 |
| Db | 101429 | CAGATGACTCCAAATTTAGGACGACGAGGCATATGGATGGGATATAAAGGGGCTGAGGC | 101488 |
| Qy | 4652 | TGAGAGCTGTCA---GAGATTCTCCAAACCCAGGTAAAGAGGAGTTTCGGTGGGGCT | 4707 |
| Db | 101489 | TGAGAGCTGTCAAGCGAGATTCTCCATCCAGGTAAAGAGGAGTTTAGCGTGGGGCT | 101548 |
| Qy | 4708 | CTTCAACCAACGACGACCTCTCCCACTAGAGGAAATCGCTTTCTGGAAGTGGGT | 4767 |
| Db | 101549 | CTTCAACCGCACGACGACCTCTCCCACTAGAGGAAAGTGTCTTCCCTGGAAGTGGGC | 101608 |
| Qy | 4768 | TCAGCCCGTCAGAGATCTGACAGAGGTGGCTTCCACGAGCTGGAAAGTTCTCAGTGGC | 4827 |
| Db | 101609 | ----- | 101640 |
| Qy | 4828 | AGGAGGTTTCCACAAGAAACAGTGCCTTCCCTTACGCTGTCTTCTCCATCTTCC | 4887 |
| Db | 101641 | AGGATGTTTC-----TACTGATGCCCTTCCCTTCCACTGCCCTCCCTC----- | 101687 |
| Qy | 4888 | TCCTGGGAGTCTCTCCCGCTTTGGTTTATCTTGGCTCTTCGCTTTCAGCAAGATTG | 4947 |
| Db | 101688 | ----- | 101731 |
| Qy | 4948 | CCCTGTGCTGCATCCATCTTCTCTACTGTCTCCGTGCTCCGTGCTGCTTCCCTTCCG | 5007 |
| Db | 101732 | CCCGTGTGCTTACTCTCCATCTGTCTACTCTCTC-----TGCTTGCCTCTTCTGTG | 101784 |
| Qy | 5008 | TGCTCTCTCTTCCACCCATTTCTCATTCTTCCCTTCTTCCCTCTCTCATTTGATTCTAT | 5067 |
| Db | 101785 | TGTTCTCTCTTCCACGTTTCTCATTCTCTGCTGCTCCCCCCCCCTTCATTTTFA- | 101843 |
| Qy | 5068 | CTTCT | 5127 |
| Db | 101844 | ----- | 101843 |
| Qy | 5128 | CCTTCT | 5187 |
| Db | 101844 | ----- | 101884 |
| Qy | 5188 | TGGGGTTC--CACCTTATGTAACAATCTTCCAGTGAGGCACACAGCTTCACTGCTGTG | 5245 |
| Db | 101895 | CAGGCTTCATACAGTCAAGTACGAATCTCCAGTACGACCAAGCTTTAATCTCTG | 101943 |
| Qy | 5246 | GGTGTCTCTTACCTTCTCTCACCCCTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 5305 |
| Db | 101944 | GGTGTCTCTTACCGACCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 102003 |
| Qy | 5306 | AGATTGGTCTCCAGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 5365 |
| Db | 102004 | -GGTCCGTCTCTCAGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 102062 |
| Qy | 5366 | CGCCACTGTGGTCTCTGTTCTCAGCTGTGTCACATCTCTCAGGATCTCTGTAAGATT | 5425 |
| Db | 102063 | CACCTCTGTGGTCT | 102122 |
| Qy | 5426 | AACCAAGGTGAGAATGTTTCCCTCTAGACAGCAGATCAAGATTCT-CCCGAAGTCAGGC | 5484 |

| | | | |
|----|--------|--|--------|
| Db | 102123 | TGAGATGCCCCAGTTTCCCTGTAGACAGCAGATCATGATTTTCCCAGAGCCAGAC | 102182 |
| Qy | 5485 | TTCCAGCCCTCTCTTCTGCCCCAGCTGCCCCGACCTCTTTAGCAAACTCAGGACACCT | 5544 |
| Db | 102183 | TTCCAGCACCAGC--CCTCTGCCAGCAACCTGGCACTCTTAGCAAACTTACGCCACCT | 102240 |
| Qy | 5545 | TACCCCATAGACC----- | 5559 |
| Db | 102241 | TCCCCACATAGACCAAGCTCTTGCAGAGGCTCTTCCAGATGACTTCGAGTTTCTGCAA | 102300 |
| Qy | 5560 | -----TCTGACAGAGAGCAGGCACTTTATCATGAGTCTCTGTTGGGAGA | 5603 |
| Db | 102301 | GGAAAGGAGAACTCTTGTGATGGGAGCAGGACCTTTTACACGAGTCTTACGCGGAGA | 102360 |
| Qy | 5604 | GCCATAGCTACGCTGTAAAGAGCGAGGAGTGGTGTAGGAAAGTCAAGACTTCA | 5663 |
| Db | 102361 | GCCATAGCTATGCAATAGCAGAGCGAGGAGTGGTGGAAAT-----TGGACTTCA | 102411 |
| Qy | 5664 | CATGAAAGCTTAGCCCCACACAGAGAAATGACAGACAGATCCCTCTCTATCTCCCCATAAGA | 5723 |
| Db | 102412 | CACAGAG-CTAGCACACACCCAGGAT---GACATATCCCTCTCTATCTCCCCATAAGA | 102466 |
| Qy | 5724 | GTTTGAG 5730 | |
| Db | 102467 | GTTTAAG 102473 | |

RESULT 5
AC119293/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

210784 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-272D8, WORKING DRAFT SEQUENCE.
AC119293
AC119293.4 GI:25072533
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 210784)
Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alabrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, I., Haviak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewari, L., Louisied, H., Lozada, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, R., Martin, R., Martinez, E.,
Manning, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Norris, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaakemele, O., Okwono, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfaunkoch, C.,

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sma's, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 210784)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23602929.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRMH
 Center clone name: CH230-272D8

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 180229 bases at least Q40
 Consensus quality: 182328 bases at least Q30
 Consensus quality: 183274 bases at least Q20
 Estimated insert size: 184593; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 210784: contig of 210784 bp in length.
 * Location/Qualifiers
 1..210784

FEATURES
 source

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-272D8"
1..1742
/misc_feature
  clone_end:Sp6
  note="wgs_end_extension"
  18139..19359
/misc_feature
  clone_end:Sp6
  note="wgs_end_extension"
  21232..22032
/misc_feature
  clone_end:Sp6
  note="clone_boundary"
  site_end:Sp6
  end_sequence:BZ243354
200775..201644
/misc_feature
  clone_end:T7
  note="wgs_end_extension"
  end_sequence:BZ243352
209704..210784
/misc_feature
  clone_end:T7
  note="wgs_end_extension"
  end_sequence:BZ243352
209704..210784
ORIGIN
Query Match      40.9%; Score 2343; DB 2; Length 210784;
Best Local Similarity 71.2%; Pred. No. 0;
Matches 4321; Conservative 0; Mismatches 795; Indels 951; Gaps 55;
QY 1 GGATCCTCGAAGTGCACCAAGGGTCTCCACCACCAAGGTGCGCTAGTCTCAATTTTCACT 60
DB 153822 GGATCCTCGAAGTGCACCAAGGGTCTCCACCACCAAGGTGCGCTAGTCTCAATTTTCACT 153763
QY 61 TTCCATGCTTGTTCCTCAATGCTGGGCTCCCGAGAGTAATTTGGAGCTTTGTTTAT 120
DB 153762 TTCCATGCTTGTTCCTCAATGCTGGGCTCCCGAGAGTAATTTGGAGCTTTGTTTAT 153704
QY 121 TTCAAAAGGGCTGATGAGGATGATCTTGTCTACCGACTCTTAAGGGTCCCGTCA 180
DB 153703 TTCAAAAGGGCTGATGAGGATGATCTTGTCTACCGACTCTTAAGGGTCCCGTCA 153644
QY 181 AGCCCTCAGACCTGGAGCTTTTGCAACAGCCCTTTAGTGGAGGAGAGTAAGCAATTT 240
DB 153643 AGCCCTCAGACCTGGAGCTTTTGCAACAGCCCTTTAGTGGAGGAGAGTAAGCAATTT 153584
QY 241 TCCTTAAAGCCAAATCCTGCTCTAGACTCTTTCTTCTGTACCTCGGTCCT-GGGCTC 299
DB 153583 TCCTTAAAGCCAAATCCTGCTCTAGACTCTTTCTTCTGTACCTCGGTCCTCGGCTC 153524
QY 300 TAGGGTGGGAGGTGGGCTTGGAAAGAGAGGTGGGAGTGGCAAAAGCCGATCCCTA 359
DB 153523 TAGGGTGGGAGGTGGGACTTG- --AAAGAGGAGGAGGAGTGGCAAAAGCCGATCCCTA 153467
QY 360 GGGCCCTGTGAAGTTCGGAGCTTCCCTGTACAGACTGGCTCATAGATCTCTCTCCAGC 419
DB 153466 GGACTCTGTGAAGTTCGAGCTTCCCGCTTCCAGTCAAGTGGCTCATAGATCTCTCTCCAGC 153407
QY 420 CAAACATAGCAAGAGTGTATACCTCTTTGTGACTTTCCCGAGGCCAGTACCTGTGAGGT 479
DB 153406 CAAACATAGCAAGAGTGTATACCTCTTTGTGACTTTCCCGAGGCCAGTACCTGTGAGGT 153347
QY 480 TGAACAGAGATTAGAGAGCTCTGAACCTGAACTCTGAAGCTCATCCACCAAGC 539
DB 153346 TGAACAGAGATTAGAGAGCTCTGAACCTGAACTCTGAAGCTCATCCAGCCATCTAC ---C 153291
QY 540 RAGCACTAGGTGCCACTGCTAGTATCTAGTATCTAGCTGATATATGACAGCTG- --GGC 597
DB 153290 ACACACCTACATGCCGGTTCTAATCAGTATCTTACACTGATATATGACAGCTGACCAAC 153231
QY 598 CACAGAGTCTGGGGTGTAGGAACCTGACCACTGACTTTTCACTGGCAAGGATGACC 657
DB 153230 GGGAGAGGCTGGAGTGTAGGAACCTGACCACTGACTTTTCACTGGCAAGGATGACC 153171

```


Db 151129 TGGCCTTCATCTCGGAGAGCAACCAATTTGGCAGAGGCACAGCAGATGGGAACGCAAGG 151070
Qy 2629 TTGAATCAGTCCCTTCAAGAATACTGCATGCAAGACACCTTAAGACCCCTGGAGAGAGGGT 2688
Db 151069 TTGAACAGGTCCTTCAAGAAATACCGGTACAGATCTAAGACCCCTGGAGAGACTGT 151010
Qy 2689 ATGCTCCTGCCCCCAACCAACCAATAGGGAGAGTAACTATCTTAAAGGGGCTGGCGACCTTG 2748
Db 151009 ATGCTCCTGCCCCCAACCAACCAATAGGGAGAGTAACTATCTTAAAGGGGCTGGCAACCTTG 150950
Qy 2749 GGGAGACACCAATTAAGAGTCTGAGAGCTGAGCCCAAGAAAACCTGAACGCCCTGTGCTCTGC 2808
Db 150949 GAGAGACACCAATTAAGAGTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 150890
Qy 2809 CCACCTCCACACTCTAGAGCTATATTTAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2866
Db 150889 CCACCTCCACACTCTAGAGCTATATTTAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 150831
Qy 2867 AGCAGGAGAGTGTTCCTGGGTGTGAGGGGTGAGGGGAAAGCCAGAGAGGGGAGTCTGG 2926
Db 150830 AGCAGGAGAGGTTCTGGGTGTGAGGGGTGGA- GGGAAAGTCAAGCAGGGGAGTCTGG 150772
Qy 2927 TTTTCTCTCTGAAACACAAATGTACTTATGATTAACAGGCAATGACCTGCTAAAGACCA 2986
Db 150771 CTTTCTCTCTGAAACACAAATGTACTTATGATTAACAGGCAATGACCTGCTAAAGACCA 150712
Qy 2987 ACATCTACGCTCTGAAAGACAGCAGCCCTGGAGGACAGGGGTGTCTCTGAGCCTTG 3046
Db 150711 ACATCTACGCTCTGAAAGACAGCAGCTCTGGAGGACAGGGGTGTCTCTGAGCCTTG 150652
Qy 3047 GGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3106
Db 150651 GGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150592
Qy 3107 GAGAGGCACTTGGGAAGGGTCACTGTCAGAGCCCTTATCCATGGAATCTGAGCCT 3166
Db 150591 GAGAGGCACTTGGGAAGGGTCACTGTCAGAGCCCTTATCCATGGAATCTGAGAC 150532
Qy 3167 GGGGCACTTGGGTAAATCTCTGGGCTGCGGAGGATCAAGAGCAGCAGCCTGATCTC 3226
Db 150531 GGGGCACTTGGGTAAATCTCTGGGCTGCGGAGGATCAAGAGCAGCAGCCTGATCTC 150472
Qy 3227 TGGCAGCTGGGAGGCGGAGGAGCAACCCCACTTATACCTTCTCTCCCTCAGGCC 3286
Db 150471 TGGCAGCTGGGAGGCGGAGGAGCAACCCCACTTATACCTTCTCTCCCTCAGGCC 150412
Qy 3287 CAGGATTAACACTCTGGCCTTCCCTTCCACCTCCCATCAGGAGTGGAGGTTGCAG 3346
Db 150411 CAGGATTAACACTCTGGCCTTCCCTTCCACCTCCCATCAGGAGTGGAGGTTGCAG 150352
Qy 3347 AGGAGGGTAAACCTACATGTCACAAACATCATGTCGACCAATATATGATGATGATGATGATGAT 3406
Db 150351 AGGAGGGTAAACCTACATGTCACAAACATCATGTCGACCAATATATGATGATGATGATGATGAT 150300
Qy 3407 GTAGAGGCAAGAAAGAAATCTGCA- - - - -GGCTTAACTGGGTTAAATGTTGTA- - - - -GT 3457
Db 150299 GTAGAGGCAAGAAAGAAATCTGCA- - - - -GGCTTAACTGGGTTAAATGTTGTA- - - - -GT 150240
Qy 3458 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3517
Db 150239 ATGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150180
Qy 3518 GTGCTGAGGTTTACAGACTGAGGTTTGTGTGTAATTTGCCCCAAGGCAAGTGGGTGAA 3577
Db 150179 GTATGTGACATTAACAGACTGAGGTTTGTGTGT- - - - -AAGTCAAGTGAAG 150135
Qy 3578 TCCCTTCACTGTTTAAAGAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3637
Db 150134 ACCCTCCATGTTTAAAGAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150075
Qy 3638 GAATGACACTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3697
Db 150074 GAATGACACTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150015

Db 755 AGCCCTGAGACCTGGAGCCTTTGAAACAGCACCTTAGGCAGAAACACAAATAAGCAATTT 814
Qy 241 TCCTTAAAGCCAAATCCCTGCTCTAGACTCTTCTCTCTAGCTCGGTCCCT-GGGCTC 299
Db 815 TCCTTAAAGCCAAATCCCTGCTCTAGACTCTTCTCTCTAGCTCGGTCCCTGGGGCTC 874
Qy 300 TAGGGTGGGAGGTGGGGCTTGGAAAGAGGTGGGAAGTGGCAAAAGCCGATCCCTA 359
Db 875 TAGGGTGGGAGGTGGGACTTGAAGAGGA---GGGHAAGTGGCAAGGCCACTCCCA 931
Qy 360 GGGCCCTGTGAAGTTGGAGCCTTCCCTGTACAGCACTGGCTCATAGATFCTCTCCAGC 419
Db 932 GGACTCTGTGAAGTTCTGAGCCTTCCAGTACAGTCTGGCTCATAGATFCTCTCCAGC 991
Qy 420 CAAACATAGCAAGATGATACCTCTTGTGACTCTCCAGGCCAGTACCTGTAGGT 479
Db 992 CAAACATAGCAAGATGATCTCTTGTGACTCTCCAGGCCAGTACCTGTCCGGT 1051
Qy 480 TGAACACAGATTTAGAGAAGCTCTGAACTCACTCACTGAACTCTGAAAGTCTATCCAAAGC 539
Db 1052 TGAACACAGATTTAGAGAAGCTCTTAACTCACTCTGGAGCTCTGAGCCCATCTAC---C 1107
Qy 540 AAGCACCTAGGTGCCACTGCTAGTTAGTATCCTAGCTGATATATGCGAGCTG---GGC 597
Db 1108 ACACACCTACATGCCGTTCTAATCAGTATCTTACACTGATATACGCAAGCTGACCCAC 1167
Qy 598 CACAGAAGTCTGGGGTGTAGAACTGACAGTGAATTTTTCAGTCGGGAAAGGTATGACC 657
Db 1168 GGGAGAAGCCTGGAGTGTAGAACTGGCAGTGAATTTTTCAGTCTGCAAGGTATGACC 1227
Qy 658 CCTCAGCAGAT---GTAGTATGTCCTTGAATCCCTAGATCCCATCCAGCAGGTCTTAAGAG 714
Db 1228 CCTTGGCAGATATGGCAGATGGGTCCCTTGAACCCCATCCAGCAGGTCTCTGAGAG 1287
Qy 715 GACATGGATGAGAGATGTAGTCACTGTGGCAATTCCTCAAAACACAGCTATCCACAGTGTCCCT 774
Db 1288 GACACAGATGACAGATGTAGTCACTGTCATTCCTCAAAACACAGCTATCCACAGTACTCT 1347
Qy 775 TGCCCTTCCACTTACGAGGAGGACAGTAACTTAGCTACTTCTTCTTCTCCCATCC 834
Db 1348 TGTCCCTTCCACTTACGATGAGGATAGCAACTTAGCTACTTCTTCTTCTCCCATCT 1407
Qy 835 TCCAGGACACACCCCTGGTCTGAGTATTCATTTCTTCTTCTTCTTCTTCTTCTTCTGAC 894
Db 1408 TCCAGGACCCAAACCCCGGTCTATGGTATTCATTTCTT-CTTCATTTCCCCACCATGAC 1466
Qy 895 TTCCATTTGCAAGGCTTTTGACCTCTGAGCTGCTGGAAGATAGAGTTTGGCCCTTAGGTG 954
Db 1467 TTCCATTTGCAAGGCTTTTACCTCTGAGTACTGGAAGATGGAGTTTGGTCTATATG 1526
Qy 955 TGGCAAGCATCTCAAGAGAAAGCAGACAAAGGGGACCAAGATTTTGAAGGATCAGGA 1014
Db 1527 TGGCAAAACCATCTCAAGAGAAAGCAGGCAACAGGGGAAACAGATTTTGTAGG-GGGAC 1585
Qy 1015 ACTAATCACTGGCGGCTGGGGTAGNAAGAGTGAAGTGAAGTTCGCTCAGCTAAG 1074
Db 1586 ACTAATCACTGGAGGGCTGGAGGTAGAAAAAAGTGA-----GCTCCAGCTAAG 1637
Qy 1075 CCAAGCTAGTCCCGAGATCTCTGCCAGCTGGGCTGCTCGGGTACTTTTAGGAATG 1134
Db 1638 CAAAGCCAGATCTTTGGGTATT-----CACAGCTGGGCTGCT-GAGGTAGTTTAGGAATG 1692
Qy 1135 TGGGTCTGAAGACAATGGATTGAAGACATCTCTTTGAGTCTCCCTCAACCCCACT 1194
Db 1693 TGGCTCTGAAGACAATGGATTTGAAGACATCTCTTTGAACCTTGGCCCTGCTGGGCT 1752
Qy 1195 ACAGACACACTGTGTGGCCAGACTCTGTGTTCAACAGCCCTCTGTGTTCTGACCACTG 1254
Db 1753 AAATCTCCCTCAACCCCACTATAGACACAGCTTGTGCAATGSCCAG-----ACTG 1802
Qy 1255 AGCTAGGCAACAGAGCATGGCCCTGTGCTGAGGATGAAGATTTGGTTTACCAATAGCAA 1314
Db 1803 AGTTAGGCAACTGGAGCATGGGCCCTGTG-TGAGGACGAGGGGTG-----CCATAGCAA 1857

Qy 1315 AAACAGCAGGGAGGAGAAACAGAAACGAAATAGGAGAGAGAGAAAGCCAGTCA 1374
Db 1858 AAACAGCAGGTGAGGAGAAACAGAAATGAAATAGGAAGAGAGCGGAACGCGC---CA 1914
Qy 1375 ATCAGATGCACTCAGAAGAGATGGGAAGCCAAACACACAGCTTTGAGCAGAGGAACAGAAA 1434
Db 1915 ATCAGATGCACTCAGAAGAGATGGGAACCAATGATAGCTTTGAACAAAGAAATAGAAG 1974
Qy 1435 AGGGAGAGA----- 1443
Db 1975 AGGAAGAGATGGCCAGATCAATTAGGGTGGACTCTCTCCATTAATTAAGATGCTCTGTGT 2034
Qy 1444 -----TTCTGGGCATAAGGAGCCACAGAAAGAGAGCCCGGCC 1484
Db 2035 CCGACTCTACATCCTCTGGGTACTGTGTCATTTGCAATCTCAGAAATAGCAGACACAGAACT 2094
Qy 1485 CCAAAGTCTCTCTTTTATACCTCATCC----- 1513
Db 2095 GTATATGGCCATTTTCACTTCTCAATCCTGAAGCTTGACACACTTGACAAATTTCTAGTT 2154
Qy 1514 ----- 1513
Db 2155 TACATTTGTGTACTCTATACAGATATGTTCTTCATGCAAGTGTGTGATGCTCCCTGAAT 2214
Qy 1514 ----- 1513
Db 2215 ACAAGTCTCTTTTATGAATCAGTGTGCAATTAATCAAGTCTGCATAGGGAGCCA 2274
Qy 1514 -----GTCTCCCAATTAAGCCCACTTCTTCTCTAGATCAGAC 1551
Db 2275 CAGGAAGAGCCAGGTCCCGAGTCTCTCTTTATGCCCCCTCTCCCTCCCTAAGCCCATTC 2334
Qy 1552 CTGAGCTCAGCAAGAGAGACCCGTAGGAGGATCACTTGGATGAAGGAGATGTGTGGAG 1611
Db 2335 CTCTTC-CTGGAACAGACTGAGCAGCAGTGGAGAAATCCGTAGAGTCTTAAAGAGATCA 2393
Qy 1612 AAGTCCAGGGCACTTAAGAGCCAGAGCCTTAAAGAGCAGAGATTAAGGTGTTCAAAG 1671
Db 2394 CAGTGGAAAGGAGATGTGTGGACAGAGCCTTAAAGAGCAAGAGATAAAGGAGCAGCAAG 2453
Qy 1672 GTGGCCAGGCTGTG--CACACAGAGGGTCTGAGACTGTGTGTAGAGCCTCAAGATAAGGA 1729
Db 2454 GCAGTACCGCTGACACACAGAGGTTCAAGACTAGCGGTAGAGCTCTCAAGATTAAGGA 2513
Qy 1730 TGATGCTCAGAAATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAGAGTGAAGGA 1789
Db 2514 TGATGCTCAGAAATGGGTGAGGAGAGAGA----- 2541
Qy 1790 GCCTGGAAACAGAGAACTCTGGAAGCCCTGGAAACGATACCATTAAGGGAGAAACCCAGGCT 1849
Db 2542 ---CGGGACTGAGACTCTGGAAGCCTGGAAACGATACTGTAAAGGGGAAACCCAGGCT 2598
Qy 1850 ACCTTTAGTGTAAATCATGAAGACAGGGAAGGAGCTGGAGAGATGAGAGGACC 1909
Db 2599 ACCTTTAGTGAAGTCAATGAAGTCAAGGAGAGGAGCTGGAGAGATGAGGAATA 2658
Qy 1910 CCGGGCAGAGACATGGAAGCAAGCAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGC 1969
Db 2659 -----AGACAAGGACNAGCCAGTTGAGTGTCTGTGAGATT----- 2695
Qy 1970 TGAAGCAGAGCCCTGGTATGAGCAACAGAACAGCAGAGGCTTAGGGTTAATCTCGAGACA 2029
Db 2696 -----CCGATATGAGCCCGGAAACAGCAGAGGCTTAGGGCTGATTTAGAGACA 2742
Qy 2030 GGGAAACAGAGGTAGACACAGAAACAGACAGAGAGCGGGGATCCAGGTTAAACAAAGGAATG 2089
Db 2743 GGGAACTGAAGTGGGCAAGGAAACAGACAGAGACGGGGGATCCAGGTTAAACAAAGGAATG 2802
Qy 2090 GTCTTCTCAGCTGTGGCCAGAGCCTCATCT-GTGTCCACATACTCTAGAAATGTTTCATC 2148
Db 2803 GTGCCCTCTCAGCTGTGGCCAGAGCCTCATCTGTGGGGTCCACTTACTCTAGAAAGTTTCATC 2862

VERSION
KEYWORDS
SOURCE
ORGANISM

AC115371.5 GI:30521556
HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 263901)
Muzny, D. Marie, Metker, M. Lee, Abranzone, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Benson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Sacoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. B., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hui, Y. K., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowls, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Loulsegod, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mathewine, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwaokemele, O., Okunolu, G., Olarnpungoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sander, W., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, C.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, R., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

2 (bases 1 to 263901)

Worley, K. C.

Direct Submission

Submitted (19-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:24954086.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center Project name: GHLJ

Center clone name: CH230-118C1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 235051 bases at least Q40

Consensus quality: 237228 bases at least Q30

Consensus quality: 238719 bases at least Q20

Estimated insert size: 244896; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 263901: contig of 263901 bp in length.

FEATURES

source

1..263901

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-118C1"

1712..259601

/note="clone boundary"

clone end:Sp6

site:EcoRI

end_sequence:BH285254"

18349..19761

/note="wgs contig"

complement(155798..262015)

/note="clone boundary"

clone end:T7

site:EcoRI

end_sequence:BH285252"

complement(261235..262017)

/note="clone boundary"

clone end:T7

site:EcoRI

end_sequence:BH285252"

ORIGIN

Query Match 28.9%; Score 1655.4; DB 2; Length 263901;
Best Local Similarity 73.0%; Pred No. 0;
Matches 2910; Conservative 0; Mismatches 451; Indels 624; Gaps 36;

QY 1833 AGGGAAGAACCCAGGCTACCTTTAGATGTAATCATGAAGACAGGAG-AAGGGAAGCT 1891

Db 263901 AGGGGAATACCCAGCTACCTTTAGATGGAAGTCAATGAAGGAGGAGGAGCT 263842

QY 1892 GGAGAGAGTAGAAGGACCCGGGGCAACATCGGAAGCAAGACCAAGCTTGAGCGC 1951

| | | | | | | | |
|----|--------|--|--------|----|--------|---|--------|
| Db | 263841 | GGAGAGGTAGAGGAATA-----AGACAAGGACAAGCCAGGTTGAGTGC | 263798 | Db | 262813 | TGACCCCGCTAAGACCCCAACACCTTCTGAAAAAAGCAGCAGTCTCTGGAGGACGGG | 262754 |
| Qy | 1952 | TCCGTGAATCAGCTGCTGAAGCAGAGCCCTGCTATGAGCACCAGAACACGACAGAGCT | 2011 | Qy | 3029 | GGTTGTCTCTGAGCCTTGGGTGCTTGTATGTTGGTGCACAAGCAGGCGCATGAGTGTGAGTAT | 3088 |
| Db | 263797 | TCTGTGAGATT-----CCGATATAGCGCCGGAAACAGCAGAGCT | 263758 | Db | 262753 | GGTTGCCGTTGGGCTTGGGTACTTGTATGTTGGTGCACAAGAGAGGCGCATGCACTGTGAGTAT | 262694 |
| Qy | 2012 | AGGGTTAATGTCCGACACGGGAAACAGAAAGGTAGACACAGGAACACAGACAGACGCGGGGAG | 2071 | Qy | 3089 | AAGGCCCCAGGAGCGTTTAGAAGAGGCACTTTGGGAAGGGTCACTCTGAGAGCCCTTAT | 3148 |
| Db | 263757 | AGGGCTGATTAGACACAGGGAATGAGGTGGGCACAGGAACACACAGAGCGGGGAT | 263698 | Db | 262693 | AAGGCCCCAGGAGGTTTAGAAGAGGCACTCGGGAAGGGTCACTCTGAGAGCCCTTAT | 262634 |
| Qy | 2072 | CCAGGTAAACAAAGGAATGCTCTTCACTGTGTGCCAGAGCGTCCATCT-GTGTCCACA | 2130 | Qy | 3149 | CCATGGAAATCTGGAGCCTGGGGCCCACTGCTGTGTTAAATCTCTGGGCTCCAGGCAATTCAA | 3208 |
| Db | 263697 | CCAGGTAAACAAAGGAATGCTCTTCACTGTGTGCCAGAGCGTCCATCTGCGGTCCACT | 263638 | Db | 262633 | CCATGGAAATCTGGAGACAGGGGCCCACTGCTGTGTTAAATCTCTGGGCTCCAGGCAATTC | 262574 |
| Qy | 2131 | TACTCTAGATGTTTCACTCAGACTGAGCGGTGGCTTGGGAGCGACCTGGAAAGATATGT | 2190 | Qy | 3209 | AGCAGCAGCTGCATCTCTCTGCGCAGCTGGGAGCGGGAAGGAGCAACCCCCACTTTATA | 3268 |
| Db | 263637 | TACTCTAGAACGTTTCACTCAGACTGAGCGGTGGCTTGGGAGCGCTGAGAAGATATGT | 263578 | Db | 262573 | AGTAGCACTGCATCTCTCTGCGCAGCTGGAGGCGGGAAGAGCAACCCCCACTTTATA | 262514 |
| Qy | 2191 | GAGAGCCAGGGGAGACAAGGGGGCCCTAGGAAGGAAGGAAGAGGGCAACCCAGGCCACACA | 2250 | Qy | 3269 | CCCTTTCTCCCTCAGCCCCAGGATTAACACCTCTGCGCTTCCCTTCCCACTCCCATC | 3328 |
| Db | 263577 | GAGAG-CGGGGGAGACTAGGGGACCGAGGAAGGAAGGAGGCGCAACCCAGGCCACACA | 263519 | Db | 262513 | CCCTTTCTCCCTCAGCCCCAGGATTAACACCTCTGCGCTTCCCTTCCCACTCCCATC | 262454 |
| Qy | 2251 | AGAGGGCAGAGCCCAAGAACTGAGTTAACTCTCTTCTGTTGATCTTCCATAGAGGCGAG | 2310 | Qy | 3329 | AGGAGTGAGGGTTTGCAGAGGGAGGTTAAACCTCATCATGTGCCAAACATCATGTGCGCAG | 3388 |
| Db | 263518 | AGAGGGCAGAGCCCAAGAACTGAGTTAACTCTCTTCTGTTGATCTTCCATAGAGGCGAG | 263462 | Db | 262453 | AGGAGTGAGGGTTTGCAGAGGGAGGGTAAGAACCTTAGAT-----AATTGTGTGCAATG | 262402 |
| Qy | 2311 | TGGGAACCTGTGACACCATCCCATGAGCGCCCTTGTCTCTGCACTGACCCCAA | 2430 | Qy | 3389 | ATATATGATCAGTATGTGTAGAGCAAGAAAGAAATCTGCA-----GGCTTAACATGG | 3442 |
| Db | 263461 | TGGGAACCTGTGACCCCAACAT-----CCCCCACTACCATATACCAAGTTTGGCCCTG | 263412 | Db | 262401 | ATATATGCACTAGTATGTGTAGAACCAAGAAAGAAATCTGGAATTAATGTGTAAGTCT | 262342 |
| Qy | 2371 | AGTGCACTTCTAGGTTTCCCTGAGGACAGAGCCTGGCCCTTGTCTCTGCACTGACCCCAA | 2430 | Qy | 3443 | GTAAATGTGTAAAA---GTCTGTGTGCATGTGTGTGTCTGACTGAAAAACGGGCACTGGCT | 3499 |
| Db | 263411 | AGTAGCACTAGGTTTCCCTGAGGAACAAAGCCCTGTTTGTCTGTTGCTTGGACCTGACCCAG | 263352 | Db | 262341 | CTGTGTGTGTGCATGTGCATGTGAATGTGTGTGTCTGACTGAAAGTGGGCATGGAT | 262282 |
| Qy | 2431 | GCTGACCCAAATGTTCTCAGTACCTTTGAAATGCCCTCAAGAGCTTGAGAACCCAGGCACTGA | 2490 | Qy | 3500 | GTGACAGCTTTCAGTCTGTGCGGTGAGGTTACAGACTGCGAGGTTGTGTGTTGTTGTTGTTGCC | 3559 |
| Db | 263351 | GCTGACCCAAATGTTCTCAGTGCCTTATCATGCCCCCAAGAGCTTGAGAACCCAGGCACTGA | 263292 | Db | 262281 | GTGACAGCTTTCAGTCTGTGTGTGATGCAATTAACAGACTGCGAGGTTGTGTGTGTGTGT | 262230 |
| Qy | 2491 | CATATTAGGCCATGGGCTAAACCTGGAGCTTGCAACAGGAGCTTCAAGTGAACCTCCAGG | 2550 | Qy | 3560 | CAAGGCAAGTGGGTGATCCCTTCCATGTTTAAAGAGATTTGGAATGATGGCTGCATCT | 3619 |
| Db | 263291 | CATATTAGGCCATGGGCTAAACCTGGAGCTTGCAACAGGAGCTTCAAGTGAACCTCCAGG | 263232 | Db | 262229 | -----AAGTCAGTGAGACCCCTCCATGGTTTAAAGAAATTTGGATGATAGCTGCATGT | 262177 |
| Qy | 2551 | GACACAGCTGCAGACAGGTGGCTTTATCCCAAGAGCAACATTTGGCATGAGTGGCT | 2610 | Qy | 3620 | CAAGGACCATGGAATATAGAAATGGAACACTTATATGTCTCTTAAGCTAAAGTACCAAGG | 3679 |
| Db | 263231 | GACACAGCTGCAGACAGGTGGCTTTATCTCTCGAGAGCAACCATTTGGCAGAGGCACA | 263172 | Db | 262176 | CAGGCAATTTGAAGTAGAATGTAGACTTTATATATGTTCTTAAGCCAAAGTAGCAGG | 262117 |
| Qy | 2611 | GCAATTTGGGAATGCAAGGTGGAATCAGGTCCCTTCAAGAATACTGCATGCAAGACCTAAG | 2670 | Qy | 3680 | TCTTTGGAGGACACTGTCTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACA | 3739 |
| Db | 263171 | GCAGATGGGAACGCAAGGTTGAAAACAGGTCCCTTCAAGAATAACCGGTACAGATCTAAG | 263112 | Db | 262116 | TCTTTGGAGGACACTGTCTAGAGATATGGGCAACAGAGCAATAGACAGTATCTGTACA | 262057 |
| Qy | 2671 | ACCCCTGGAGAGGGGTATGCTCTCTGCCCCACCCACCATTAAGGGGAGTGAACATCTCT | 2730 | Qy | 3740 | GAGTAAGGAGAGAGAGAGGGGGTGTAGAAATCTCTTACTACTCAAGGGGAACCTGATCG | 3799 |
| Db | 263111 | ACCCCTGGAGAGGACTGTATGCTCTCTGCCACCCACCCACCATTAAGGGGAAGTGAAGTATCCT | 263052 | Db | 262056 | GAGTAA--AGAGAGAGGAGTGAAGTATAGAAATCTCTTACTACTCAAGGGGAACCTGATCTA | 261999 |
| Qy | 2731 | AGGGGGCTGGCCACCTTGGGAGACACCACTTACTGAGAGTGTGAGCCCAAGAAAACCT | 2790 | Qy | 3800 | TGCACCTGCCAAAGTGGATGCTCTCCCTAGACATCATGACTTTGTCTCTGGGAGGCAGCA | 3859 |
| Db | 263051 | AGGGGGCTGGCCAACTTTGGAGAGACACCACTTCAOGAGAGTGTCTACTAGAAAACCT | 262992 | Db | 261998 | TGCACCTGCCAAATGGAATGCCCTCCCTGGACATCATGACTTTGTCTCTGGGAGGCAGCA | 261939 |
| Qy | 2791 | GACCCGCTGTCTCTGCCACCTCCACACTCTAGAGCTTATATT--GAGAGGTGACAGTA | 2848 | Qy | 3860 | CTGTGGAACTTTCAGGTCTGAGAGAGTGAAGGCTCCCTCAGCCCTGGAAGCTATGAGATA | 3919 |
| Db | 262991 | GACCCGCTGTCTCTACCCACCTCCACACTCTAGAGCTTATATT--GAGAGGTGACAGTA | 262932 | Db | 261938 | CTGTGGAACTTTCAGGTCT--GAGAGTGAAGGCAACCCCTCAGCCCTGGAAGCTGTGAGATA | 261881 |
| Qy | 2849 | GATAGGTGGAGCTGTGATGAGGAGAGTGTCTCTGGGTGTGAGGTGTAGGGGAAAGC | 2908 | Qy | 3920 | GCCAGGTTGAAA--GGGGGAAGGAGAGCTGGGATGGGAGCTGTGTGTGGAGGAG | 3977 |
| Db | 262931 | GAT--GGGGGAGAGCTGGGAGCAGGAGAGCGTCTCTGGGTGTGGGGTGGGA--GGGAAAGT | 262874 | Db | 261880 | GCTAGGTTGTAAGAGAGGGAAGGGGGAGGCTTGAATGGGAGCTTGTGTGTGGAGACAG | 261821 |
| Qy | 2909 | CAGAGCAGGGAGTCTGGCTTTGTCTCTGGAACAACTGCTACTTGTATTATAACAGGCA | 2968 | Qy | 3978 | GGGACAGATATTAAGCCCTTGGAAAGAGAGTGAACCCCTTACCCAGT--TGTTCAACTCACCT | 4036 |
| Db | 262873 | CAGAGCAGGGAGTCTGGCTTTGTCTCTGGAACAACTGCTACTTGTATTATAACATGCA | 262814 | Db | 261820 | GGGACAAATATTAGCCCGTGAAGAGAGTGAACCCCTTACCCAGTGTGTCTCACTCAGCT | 261761 |
| Qy | 2969 | TGACCTGTCTAAAGCCCAACATCTACGACCTCTGAAAAAGACAGCAGCCCTGGAGGACAGG | 3028 | Qy | 4037 | TCAGATTTAAAAATTAACCTGAGGTAAAGGGCC---TGGGTAGGGGAGGTGTGTGAGAGCTC | 4093 |
| Db | | | | Db | 261760 | TCAGATTTAAAAATTAACCTAAGGTAAAGGGCCATGTGTGGGTAGGGAGGTGTGTGAGAGCGGTC | 261701 |

```
QY 4094 CTGCTCTCCTCGCATGCCCTGAGGCGCTTTGGGAGGAGGAATGTCGCCAAGACTAA 4153
DB 261700 CTGCTCTCCTCTATCTGCCCCATCGGCCCTTTGGGAGGAGGAATGTCGCCAAGACTAA 261641
QY 4154 AAAAGGCCATGAGCGAGAGGGCGAGGCAACAGACCTTTTCATGGGCAAACTTTGGGG 4213
DB 261640 AAAAGGCCCTGAGCGAGAGGGCGAGGCGAGACCTTTTCATGGGCAAACTTCAGG 261581
QY 4214 CCGGTAGTGATCGATTGACAGAACTCGCCAAATCGATACCTTCTTCTTAACGACAG 4273
DB 261580 C-----TGCTGTCTCTCTGTCACTCCAGAGCC 261553
QY 4274 GAGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTCAGAGTCAGTTTAC 4333
DB 261552 AAGGATCAAGAGAGAGAGCGCAGGACAGGAGGATGGAGGAG----- 261507
QY 4334 CACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATA 4393
DB 261506 ----- 261507
QY 4394 GAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGA 4453
DB 261506 ----- 261507
QY 4454 GTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCA 4513
DB 261506 ----- 261507
QY 4514 GTGATAGAGAAAAGTGAAAGTCGAGTTTACCCTATCCTATCAGTGATAGAGAAAAGTGAA 4573
DB 261506 ----- 261507
QY 4574 AGTCAGCTCGGTAACAGCAGAGGACTCCAAATTTAGGAGCAGGATATGGATGGGAT 4633
DB 261506 -----GGTCCAGCAGATGACTCCAAATTTAGGAGCAGGACGCTGGAATGAGCT 261457
QY 4634 ATAAAGGGCTGGACACTGAGAGCTGTCAG-----AGATTTCTCAACCCAGGTAAGAG 4689
DB 261456 ATAAAGGGGCTGGAGCGCTGAGAGCTGTCAGACCGAGATTTCTCCATCCAGGTAAGAG 261397
QY 4690 GAGTTTCGGTGGGGGCTCTTCAACCAACACAGAGCTCTCCGCCACTAGAGGAAACTGC 4749
DB 261396 GAGTTTAGGTGGGGGCTCTCAACCGACACAGAGCTGTCCCACTAGAGGAAAGTGT 261337
QY 4750 CTTTCTGGAAGTGGGGTTCAGGCCGCTCAGAGATCTGACAGGGTGGCTTCCACAGCC 4809
DB 261336 CTTTCTGGAAGTGGGGCT-----CCTCCACCGGCC 261305
QY 4810 TGGGAAGTTCTAGTGGCAGGAGGTTTCCAGAGAAACACTGGATGCCCTTCCCTTAGG 4869
DB 261304 TGGGAAGTTCTGGTGGCAGGAGTGTCTA-----CTGGATGCCCTTCCCTTCCA 261254
QY 4870 CTGCTCTTCCATCTTCTCTCGGGATGCTCTCCCGTCTTGGTTTATCTTGGCTCTT 4929
DB 261253 CTGCTCTTCCCTC-----CCTGTCTGATTAATCTTGGCTCTT 261214
QY 4930 CTTCTTACAGAAATTTGCCCTGTGCTGTGTCACACTCTTCTCTACTGTCTCGGTGCC 4989
DB 261213 AGTGTTTCAGAAAGATTGGCCGGTGTGTCTACTCACTGTCTCTACTCTCTC----- 261160
QY 4990 TTGCTTGGCTTCTTGGGTGCTTCTTTCACCGGATTTCTCACTTCACTTCTTCTGCC 5049
DB 261159 -TGCCTTGGCTTCTTGGGTGCTTCTTTCACCGGATTTCTCACTTCACTTCTTCTGCC 261101
QY 5050 CTTCTCATTTGATTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5109
DB 261100 CCCCCCTTCATTTTATCTTCTTCTTCTT----- 261070
QY 5110 TTCTTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5169
DB 261069 -----CTGTGTCTAG 261061
```

```
QY 5170 AGTGTGAGAAATCACACCTGGGGTTCACACC-----TTATGTAAACAATCTTCCAGTGAGCC 5227
DB 261060 AATGTGGGAATCAAAACCCAGGGCTTCATACAGTCAAGTAAGCAATCTCCAGTGAGCC 261001
QY 5228 ACAGCTTCAGTGTCTGCTGGGTGCTCTCTTACCTTCTCACCCCTGGCTTGTCTGTCTTCC 5287
DB 261000 AAAGCTTTAATCTCT-CTGGGTGCTGTCTTACCGACCTCACTCCCTGTCTTGTCTGTCTCC 260942
QY 5288 ATCTGTGTCAAGATCTCTAGATTGCTTCCAGAGCTCTGTACTCTCTTCTTCTGCTGTCT 5347
DB 260941 GTCTAGTCAAGATCTCT-GGTCCGTCTCTCAGTCTTCTGTACTCTCTCTCCCTGCTGTCT 260883
QY 5348 COTCTCTCTGTCAGCTCGCCACTGTGTGTCTCTCTTCTTCTGTCAGCTGTGTGTCCACATCTTTC 5407
DB 260882 CTTCTCTCTGTCAGCTCGCCACTGTGTGTCTCTTCTTCTGTCAGCTGTGTGTCCAAATCTCTCT 260823
QY 5408 AGGATTTCTGAAAAGTTAAACAGGTGAGAATGTTTCCCTGTAGACAGCAGATCACGAT 5467
DB 260822 GTGAAAAGATTAAACCGGTGAGAATGCCCCAGTTTCCCTGTAGACAGCAGATCATGAT 260763
QY 5468 TCT-CCCGAAGTCAAGCTTCCAGCCCTCTCTTCTTCTTCTGTCAGCTGTGTGTCCGACCTCTTCA 5526
DB 260762 TTTCCCGAAGCAGACTTCCAGCACCCTG-CTCTGTCAGCAGCAACCTGTCACCTCTTCA 260705
QY 5527 GCAAAACCTCAGCACCCTTACCCACATAGACC----- 5559
DB 260704 GCAAAACCTCAGCACCCTTCCCCACATAGACCAAGTCTTGACAGAGAGCTCTCTTCTTCA 260645
QY 5560 -----TCTGACAGAGAAAGCAGGACCTTTTACA 5585
DB 260644 GACTTCGAGTTTCTGCAAGGAAAGAGAGAACTCTTTGTGTGATGGGAGAGGACCTTTACA 260585
QY 5586 TGGAGTCTGTGGGAGAGCCATAGGCTACGGTGTAAAGAGGAGGAGGAGTGTGTGTGTGT 5645
DB 260584 CGGAGTCTGTGGGAGAGCCATAGGCTATGTCATAGCAGAGGAGGAGGAGTGTGTGTGT 260525
QY 5646 AGGAAAGTCAGGACTTTCATAGAGCTTACCCACACAGCAACAGCAAGATCCCTCCCT 5705
DB 260524 -----TGAGCTTTCACAGAAAG-CTAGCACACACAGGAAT-----GACATATCCCT 260479
QY 5706 CCTATCTCCCCATAAGAGTTTGAAG 5730
DB 260478 CCTATCTCCCCATAAGAGTTTGAAG 260454
```

```
RESULT 8
HAMSHCA
LOCUS
DEFINITION Mesocricetus auratus alpha-cardiac myosin heavy chain gene exons
1-39, complete cds.
ACCESSION L15351
VERSION L15351.1 GI:402373
KEYWORDS alpha-cardiac myosin heavy chain.
SOURCE Mesocricetus auratus (golden hamster)
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE
1 (bases 1 to 32415)
AUTHORS Wang, R., Sole, M.J., Cukerman, E. and Liew, C.C.
TITLE Characterization and nucleotide sequence of the cardiac
alpha-myosin heavy chain gene from Syrian hamster
J. Mol. Cell. Cardiol. 26 (9), 1155-1165 (1994)
JOURNAL 95115033
MEDLINE 7815459
PUBMED 7815459
COMMENT Original source text: Mesocricetus auratus (strain F1B) male adult
liver DNA.
FEATURES
source
Location/Qualifiers
1..32415
/organism="Mesocricetus auratus"
/mol_type="genomic DNA"
/strain="F1B"
/db_xref="taxon:10036"
```


| | | | | | |
|--------|----------------------------|----|------|---|------|
| exon | /number=20 17338..17593 | Qy | 822 | TTCTCTCCCATCTCCAGGACACACCCCTGCTGAGTATTCATTT-----CTTCC | 875 |
| intron | /number=21 17594..17839 | Db | 585 | TTCTTTTCTCCCACTCTGTGTCCACCCAGGACCAACCCCTTGCTCGTATCTCAG | 644 |
| exon | /number=21 17840..18082 | Qy | 876 | TTCAAGTCCCTCTGTGACTTCCATTTGCAAGGCTTTTGACCTCTGACCTGCTGGAAGA | 935 |
| intron | /number=22 18083..18254 | Db | 645 | TTCTTTCTCCCTCATTTTCCCTCCATCTGCAAGGCTTTTGACCTCTGACGACGCTGGAAGA | 704 |
| exon | /number=22 18255..18431 | Qy | 936 | TAGAGTTTGGCCCTAGGTGTGCAAGCCATCTCAAGAGAAAGCAGACAAACAGGGGACCA | 995 |
| intron | /number=23 18432..18949 | Db | 705 | TGGGATTTTGGCTTTATGCGGCAAGCCATCTCAAGAGGAAGCAGGACCAACAGGAAACCA | 764 |
| exon | /number=23 18950..19095 | Qy | 996 | GATTTTGGAAAGGATCAGGAATCTACTGCGGCGCTTGGGGGTAGAAAAAAGAGTGTAG | 1055 |
| intron | /number=24 19096..19294 | Db | 765 | GATTTTGGAAAGGAGGAGGACTAAATCATTTAGAGGGCCATGAGGTAGAGAGAGAAACAAG | 824 |
| exon | /number=25 19295..19385 | Qy | 1056 | TGAGTCCGCTCCAGCTAAGCCAGCTAGTCCCCGAGATCTCTGCCACAGCTGGGCTGCT | 1115 |
| intron | /number=25 19386..21100 | Db | 825 | CAAGTCTCTGACGCTAAGCAAGGCTAGTTCCCTCGGCATTTCTGCTCA--GGGCTGCT | 881 |
| exon | /number=25 21101..21490 | Qy | 1116 | CGGGTAGCTTTAGGAATGTGGTCTGAAAGACAATGGGATTTGGAGAGACATCTCTT---- | 1171 |
| intron | /number=26 21491..21819 | Db | 882 | AAGGGCAGGGTTAGGAATGTGGTTTCGAGAGACAATGGGATTTTAAATCCATCTCTGGGA | 941 |
| exon | /number=26 21820..21946 | Qy | 1172 | -----TGAGTCTCCCTCAACCCCACTTACAGACACATCTCGTGTCT | 1212 |
| intron | /number=27 21947..22036 | Db | 942 | ACCTGGTTCAGCCCTGGGCGGAAATCTCCCTCAACCATACCTATAGACACATTTGTGCAC | 1001 |
| | | Qy | 1213 | GGCCAGACTCTCTGTTCAACAGCCCTCTGTGTTCTGACCACCTAGCTAGGCAACACAGACA | 1272 |
| | | Db | 1002 | GGCCGACTCTCTGTTCAACACTCTCCCTGGGCTGTGATCACTGAGAGGCACACAGACA | 1061 |
| | | Qy | 1273 | TGGGCCCTGTGTGAGGATGAAGATGTGTTACCAATAGCAAAAACAGCAGGGGGAG | 1332 |
| | | Db | 1062 | TGGGCCCTGTGTGAGTGTAGAGG---GGTTGCCAACAGCAAAAGCAGCAGGGGGGG | 1117 |
| | | Qy | 1333 | AACAGAACCAATAAAGGAGGAGAAAGGCCAGTCAATCAGATGCAGTCAGAAG | 1392 |
| | | Db | 1118 | AACAGAAAGCCATAAGGAAGAA----- | 1142 |
| | | Qy | 1393 | AGATGGGAGCCCAACACACAGCTTGAGCAGAGGAAACAGAAAGGAGAGATTTCTGGGCA | 1452 |
| | | Db | 1143 | --AAGGAATGCTTACTCAGGTTAGACTGGGAGCGGGATGCTGGGCA----- | 1192 |
| | | Qy | 1453 | TAAAGAGCCACAGAAAGAGAGCCAGGCCCCCAAGTCTCTCTTTATACCTCATCC | 1512 |
| | | Db | 1193 | ----GAGGCAACAGAAAGAGAGAGCCCAAGCCCACTTCTTTTACACCTCAATCCC | 1248 |
| | | Qy | 1513 | CGTCTCCCAATTAAAGCCACTTCTTCTTAGATCAGACCTGAGCTGACGCAAGAGACC | 1572 |
| | | Db | 1249 | TTCTTTCTCTTAAACCATTCCTCTCCCTAAGAACAGACTTGAACAGCAAGAGAGATC | 1308 |
| | | Qy | 1573 | CGTAG-----GGAGGATCACACTGGATGAAGGAGATGTGGAGAGTCCAGGGCAACC | 1626 |
| | | Db | 1309 | AGTAGTCAGAGAGGTACACCCGCAAGAGGTGATGTGTGAGAGGCCCAAGAAAGCT | 1368 |
| | | Qy | 1627 | TAAAGCCAGAGCCCTTAAAGAGCAAGAGATAAGGTGCTTCAAGGTGGCCAGG--CTGTG | 1685 |
| | | Db | 1369 | AAGAGCCAGAGCCCTTAAAGAGCAAGAGCTAAAGATGCCACAAAGGTAGCCAGGACTGG | 1428 |
| | | Qy | 1686 | CACACAGAGGTCGAGGACTGGTGTAGAGCTCAAGATTAAGATGATGCTCAGATGGG | 1745 |
| | | Db | 1429 | CCTGAACAGAGTCAAGGACGCGGTAGAGCCCC-----AGCAGAGGACGCTAAGAAATGCA | 1483 |
| | | Qy | 1746 | CGGGGGGGGGATTTCTGGGGGGGAGAGAGAGGTGAGAGAGGCTGGAAGAGCTGGAACAGAGAT | 1805 |
| | | Db | 1484 | CGAG-----GAGAGACAATGGGAAGGGGCTTGGAAACGGAGAT | 1521 |
| | | Qy | 1806 | CTGGAAGCGCTGGAAACGATACATAAAGGGAAGAACCCAGGCTACCTTTTAGATGTAAAT | 1865 |
| | | Db | 1522 | CTGAGGACCGGAGGTGACAC--TGCAAGGGAAGAACTCAGGCTACCTTT--CATGTAAAT | 1578 |
| | | Qy | 1866 | CATGAAGACAGGGAGAGGGGAAGCTGGAGAGAGTAGAAGGACCCCGGGGGCAAGCATGG | 1925 |

Db 1579 CACGAA-----GCCAAGGAAAGGAGGCTGGAGAGGAAGAACCCGGGGCAAGACATGG 1631
QY 1926 AAGCAAGACAAAGCAGGTTGAGCGCTCGGTGAATCAGCCCTGTGTAAGGCGAGCCCTG 1985
Db 1632 -----TAGCCAGGCGAGCACTCTGTGAGATCTGTGTCAGAA--GCAGAGCCCTG 1680
QY 1986 GTATGAGCACCAAGACAGCAGAGGCTAGGGTTAATGTCAGACAGGGAACAGAGGTAGA 2045
Db 1681 GTGCGAGCGCCAGAGAGGAGGTC-----TGATTCGGGACGGAAACAGAGGTAGA 1731
QY 2046 CACAGGAACAGACAGAGAGC--GGGAGAGCAGGTAAACAAAGGAAT----- 2088
Db 1732 CCCTGGAACAGACCAAAACGCGAGGAAGCCAGGGAACAAAAGAGCGACTGTGCCCTCAAAG 1791
QY 2089 -----GGTCTCTTCTCACTGTGGCCAGAGCGTCCATCTGTGTCCACATAC 2133
Db 1792 CCAGCTTGGCCCTTGGGTCACTCTTACCTGTGACAGAGCGTCCATCTGTCTTCCACATAC 1851
QY 2134 TCTAGAATGTTTCATCAGACTGCAGGCTGGCTTGGGAGGCAGCTGGAAAGAGTATGTGAG 2193
Db 1852 TCTAGAACTTTTCATCAGACTGCCGGGCTG-----GGAGGCTCTGAAGAGTATGTGAG 1903
QY 2194 AGCCAGGGGAGACAAGGGGGCTTAGGAAAGGAAGAGGGCAAAACCRAGGCCACACAAGA 2253
Db 1904 AGCCAGGAGAGAGAGGAGGCTAGGAAAGAAAGAAAGGCAAAACAGGCCACACA-- 1960
QY 2254 GGGCAGAGCCAGAACTGAGTTAACTCTTCTGTGTGATCTTCCATAGAGGCGAGTGG 2313
Db 1961 -----AAAGCCACAGAGCTGAGTTAACTCTTCTCTGT-----CTCCACAGGAGACAATGG 2010
QY 2314 GAACCTCTGTGACACACNTCCCATGAGCCCCCACTACCCATACCAAGTTGGCCCTGAGT 2373
Db 2011 AGACTCCGGTGCACAACTCTCCCAATCCCAATCCCACTACCTATACCAACTTTGTCTGTGT 2070
QY 2374 GGCACTTCTAGGTTCCCTCAGGACAGAGCTGCGCTTTGTCTTGGACCTGACCCCAAGCT 2433
Db 2071 GGCAACAGGGGTTCCCTGAGGACAGATGATGTCTTGTCACTTGGACTTGGACCCAGGCT 2130
QY 2434 GACCAATGTTCTAGTACCTTTGAATGCCCTCAAGAGCTTGAGAACAGAGCAGTGACAT 2493
Db 2131 GACCAATGTTCTAGTCCCTTATCATGTCTTTCAGGACCTTGAAGAACAGGCGAGTGACAT 2190
QY 2494 ATTAGGCCATGGCTTAAACCTGGAGCTTGACACAGGAGGCTCAAGTGACCTCCAGGGAC 2553
Db 2191 ATTAGGCCACAGGCTAACCTGTGACTTGCACAAGGTGACCT-----TCCAGGGAC 2241
QY 2554 ACAGCTGCAGACAGGTGGCTTTATCCCAAGAGCAACATTTGGCATAGGTGGCTGCA 2613
Db 2242 CTAGCTGCAGACAGGTGGCTTGCACTCTGAGAACATCATTTGGCATAGTCACTGCA 2301
QY 2614 AATGGGAATGCAAGGTTGAAATCAGGTCCCTTCAAGA-----ATACTGCATGCAAGACCTAA 2669
Db 2302 GATGGGAATACAAAGTTGGCTCAGGTCCCTTCAAGAGAACTAACTGTATACATGAGCTAA 2361
QY 2670 G-----ACCCCTGGAGAGGGGTATGCTCTGCGCCCAACCCACCAATAGGGGAGT 2720
Db 2362 GTGTGTGCGTGCCTGGAGAGGTGTGTATGCCCCCTGAGCCCAAAATTTGGCAGGGAGGT 2421
QY 2721 GAACTATCTTAGGGGCTGGCACTCTGGGAGACACCAATTACTAGAGGTGCTGAGCC 2780
Db 2422 GTCTTA-----GGGGCTGGTGACCTTGGGGGACACCG--TCACGTGAGATTGCTGTACC 2474
QY 2781 CAGAAAAATGACCGCCCTGTGTCTGCCCACTCCACACTCTAGAGCTATATTGAGAGG 2840
Db 2475 CAGAGAAATTTGACCGCTGTGTCTTACCTACCCACCTCCACTCTAGAGCTATATTGAGAGG 2534
QY 2841 TCACAGTACATAGGGTGGAGCTGGTAGCAGGAGAGTCTTCTGGGTGTGAGGTGTAG 2900
Db 2535 TGGCAGTAGTAGGGTGG-----GTGAGTCTCTGGGTATGACGTGTAG 2578
QY 2901 GGGAAAGCCAGAGCGGGGAGTCTGGCTTTGTCTCCTGAAACAAATGTCTACTTATTTAT 2960

Db 2579 GGG--AAGCCAGAGAAGGGGAGTCTGACTTTGTCCCTGGAACACAATGTCTCTTAGTTAT 2637
QY 2961 AACAGGATGACCTGCTAAAGAACCCAAACATCTAGACCTCTGAAAGACAGCAGCCCTGG 3020
Db 2638 AACAGGATGGCCTGTTAAGGACCTAACACCTACGGCCTCTGAGA--AATGGCTCTGTA 2695
QY 3021 AGGACAGGGGTTGTCTCTTGAGCCTTTGGGTGCTTGATGTTGTCACAAAAGAGGCGCATGAGT 3080
Db 2696 AGGACAGGGAACCTCTAGTGGCCGGGTCTTGAAGGTGCCACAGAGAAATGTTGTATA 2755
QY 3081 GTGAGTATAAGGCCCCCAGGAGCGTTA--GAGAAAGGCACTTGGGAAGGGGTGAGTCTGAG 3139
Db 2756 AGCCCAAGGAGGTTAGAGAGGGGTACCAAGAGTTCTTACTCTGGGAAGGAGAGTCTGGG 2815
QY 3140 AGCCCTTATCATGGAATCTGGAGCCTGGGCGCAACTGGTGTAAATCTCTCG----- 3191
Db 2816 AGCCCGCACCACTGGAATTCGGAGCCTGGGCGCAATGGGTATAAATCTCTGAGGTGCCA 2875
QY 3192 ----GCCTGCGCAGGCATTCAAAAGCAGCACCTTGCATCTCTGCGACGCTTGGGAGCGGAA 3247
Db 2876 GTCCCCCAGCAGCCATCCCAAGCGGCACCTGCACTCTCTGCGACCTTGAAGAGCGAA 2935
QY 3248 GGGAGCAACCCCCCACTTATACCTTTCTCTCAGCCCCCAGGATTAACACTCTGGCC-- 3306
Db 2936 GGGAGCAACCCCCCACTTACACCCCTCTCTCTCAGCCCTAGGATTAACACTCTGGCCT 2995
QY 3307 TTCCCTTCCCTCCACTCCCATCAGAGGTGGAGGTTGCAGAGGAGGGTAAAAACCTACA 3366
Db 2996 TTCCACCTCTCATCCCCCATCAGAGGTGGAGGTTACAGAGGAGGGTAAAAACCTGTG 3055
QY 3367 TGTCAAACATCATGTTGTCAGCATATATGGAATCAGTATGTGTAGAGGCAAGAAAGGAAT 3426
Db 3056 TGTCAAACATGTTGTTGCA--GATATATGACACAGTACATGTAGACGCAAAAAGAAAT 3114
QY 3427 CTCAGGCTTAACTGGGTTAATGTAAAGTC--TGTGTGATGTGTGTGTCTGACTGA 3485
Db 3115 CTGCGACTTATGCTGGGTTAATGTGTGAAGCGTGTGTGTGTGTGTGTGTGTGTGTGT 3174
QY 3486 AAAAGGCGATGGCTGTGCGAGCTGTTCAGTTCTGTGCGTGAGGTTACACAGCTCAGGTTT 3545
Db 3175 GTGTGTGTGTGTC--GCACAGCTGTTCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3233
QY 3546 GTGTGTAATTTGCCCAAGGCCAAAGT--GGTCAATCCCTTCCATGTTTAAAGAGATTGGA 3604
Db 3234 GTTTGTAATTTGCCCAAGGCCAGAGTGGGGTGAATCCCTCCCGTGTGTTAAAGACATTGA 3293
QY 3605 TGATGGCCTGATCTCAAGAGCAATGGAAATAGAAATAGAACCTCTATATGTGTCTTAA 3664
Db 3294 TGATGG--CGGCATGTCCGGGCAAT--CAGAGTAAGATAGATACTCTATATGTGTCCCCAA 3351
QY 3665 GCTAAGGTAGCAAGTCTTTTGGAGACACCTGTCTTAGAGATGTGGGCAACAGAGACTACA 3724
Db 3352 ACCAAGGTAGCAAGGCTTTGGAGGATACCTGTCTTAGAGATGTGGGCAACAGAGACCATA 3411
QY 3725 GACAGTATCTGTACAGAGTAAGGAGAG-----AGAGGAGGGGT 3763
Db 3412 GACAGATTTATACAGATTAAGGAGGGGAGGAGAGAGAGAGAGAGAGAGAGAGGGGG 3471
QY 3764 GTAGAAATCTCTTACTATCAAGGGAATCTGAGTCGTGCACTGCAAAAGTGTGATCTCTC 3823
Db 3472 AGTAATTTTTTTTACTAAACAAAGAAACTGAGTCATGATCTGAACAATGATGTCTCTC 3531
QY 3824 CCTAGACATCATGATTTGTCTCTGGGAGCCAGCACTGTGGAATCTCAGGCTCAGAGA 3883
Db 3532 CTTGACATCTTGAATTTGATCTCTGGGATCTTAGCACTGTGGAACTCAGGCTCAGAGG 3591
QY 3884 GTAGAGGCTCCCTCAGCCTGAAGCTATGAGATAGCAGGCTTGAAGGGGAGGGA 3943
Db 3592 GTAGACGATGCCCTCAGCCTGAAGCTGTGACATAGCCAGGAGGAGCGGGGA----- 3646
QY 3944 GAGCCTGGATGGGAGCTTGTGTGTGAGGCGAGGGAAGATATTAAGCTTGAAGAGA 4003
Db 3647 -----GGGAGCCTGTGAGTTGAAGGCTGGGGATAAGTATTAGGCTGTGAAGAGA 3695


```
Db 849 CGTCTTGGTTATCTTTGGCTCTTCGCTTTAGCAAGATTGGCCCTGTGTCACCTCA 908
QY 4967 TCTTTCTCTACTGTCGCGTCCCTGCTTGCCTCTCTGCTGCTGCTCTCTTTCCACCCA 5026
Db 909 TCTTTCTCTACTGTCGCGTCCCTGCTTGCCTCTCTGCTGCTGCTCTCTTTCCACCCA 968
QY 5027 TTTCTCACTTCAGCTTTTCTCCCTTCTCATTTGTATTATCATCTCTCTCTCTCTCTCT 5086
Db 969 TTTCTCACTTCAGCTTTTCTCCCTTCTCATTTGTATTATCATCTCTCTCTCTCTCTCT 1028
QY 5087 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5146
Db 1029 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1088
QY 5147 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5206
Db 1089 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1148
QY 5207 TAAACAATCTTCAGTGAGCCACAGCTTCAGTGCTCTCTGCTGCTCTCTTACCTTCTCA 5266
Db 1149 TAAACAATCTTCAGTGAGCCACAGCTTCAGTGCTCTCTGCTGCTCTCTTACCTTCTCA 1208
QY 5267 CCCCCTGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5326
Db 1209 CCCCCTGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1268
QY 5327 CTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5386
Db 1269 CTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1328
QY 5387 CAGCTGTGGTCCACATCTTTCAGGATCTCTGAAAAAGTTAAACCAGCTGAGATGTTTCC 5446
Db 1329 CAGCTGTGGTCCACATCTTTCAGGATCTCTGAAAAAGTTAAACCAGCTGAGATGTTTCC 1388
QY 5447 CTGTAGACAGAGATCAGATCTCTCCGGAAGTCAGGCTTCAGGCTCTCTTTCTCTCTG 5506
Db 1389 CTGTAGACAGAGATCAGATCTCTCCGGAAGTCAGGCTTCAGGCTCTCTTTCTCTCTG 1448
QY 5507 CCAGCTGCCCCGACCTTTAGCAAACTCAGGACACCTTACCCACATAGACCTCTGACA 5566
Db 1449 CCAGCTGCCCCGACCTTTAGCAAACTCAGGACACCTTACCCACATAGACCTCTGACA 1508
QY 5567 GAGAAGCAGGCACTTTACATGAGTCTCTGCTGGAGAGCCATAGGCTACGGTGTAAGA 5626
Db 1509 GAGAAGCAGGCACTTTACATGAGTCTCTGCTGGAGAGCCATAGGCTACGGTGTAAGA 1568
QY 5627 GGCAGGGAAGTGTGTAGGAAAGTCAGGACTTCACATAGAGCTTAGCCACACCAG 5686
Db 1569 GGCAGGGAAGTGTGTAGGAAAGTCAGGACTTCACATAGAGCTTAGCCACACCAG 1628
QY 5687 AATGACAGACAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5731
Db 1629 AATGACAGACAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1673
```

RESULT 12

```
MUSCMHCG
LOCUS MUS musculus cardiac myosin heavy chain gene, 5' end.
DEFINITION M62404.1 GI:192609
ACCESSION M62404.1
VERSION 1 (bases 1 to 2594)
KEYWORDS cardiac myosin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Gulick,J., Subramaniam,A., Neumann,J. and Robbins,J.
AUTHORS Isolation and characterization of the mouse cardiac myosin heavy
TITLE chain genes
JOURNAL J. Biol. Chem. 266 (14), 9180-9185 (1991)
MEDLINE 91225025
PUBMED 2026617
```

```
COMMENT Original source text: Mus musculus DNA.
FEATURES
  Location/Qualifiers
    1..2594
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      join(603..623,1305..1373,1665..1880)
      /product="cardiac myosin heavy chain"
      603..623
      /number=1
      624..1304
      /number=1
      1305..1373
      /number=2
      1374..1664
      /number=2
      1665..1880
      /number=3
      1680..1880
      /codon_start=1
      /product="cardiac myosin heavy chain"
      /protein_id="AA37424.1"
      /db_xref="GI:192610"
      /translation="MTDAQMADFGAAQYLRKSEKERLEAOTRPFDIRTECFVPDDKE
      EYVAKVVSREGGKVTAETENGK"
      1881..>2594
      /number=3
  intron
ORIGIN
Query Match 22.3%; Score 1279.8; DB 10; Length 2594;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;
QY 3767 GAAATCTCTTACTATCAAAGGGAATACTGAGTGTGACCTGCAAGTGAAGTGTCTCCCT 3826
Db 1 GAAATCTCTTACTATCAAAGGGAATACTGAGTGTGACCTGCAAGTGAAGTGTCTCCCT 60
QY 3827 AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGAACTTCAGGTCTGAGACAGTA 3886
Db 61 AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACCTTCAGGTCTGAGACAGTA 120
QY 3887 GGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCAGGTTGAAAGGGGAGGAGAG 3946
Db 121 GGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCAGGTTGAAAGGGGAGGAGAG 180
QY 3947 CCTGGATGGGAGCTTGTGTGTGGAGGAGGAGCAGATATTAAGCCTGGAAGAGAG 4006
Db 181 CCTGGATGGGAGCTTGTGTGTGGAGGAGGAGCAGATATTAAGCCTGGAAGAGAG 240
QY 4007 TGACCCCTTACCCAGTTGTTCAACTCACCTTCAGATTAAAAATACTGAGGTAAAGGCCT 4066
Db 241 TGACCCCTTACCCAGTTGTTCAACTCACCTTCAGATTAAAAATACTGAGGTAAAGGCCT 300
QY 4067 GGGTAGGGAGGTGTGTGAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4126
Db 301 GGGTAGGGAGGTGTGTGAGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 4127 GGGAGAGGAATGTGCCCAAGGACTTAAAAAAGGCCATGGAGCCAGAGGGGAGGGGCAA 4186
Db 361 GGGAGAGGAATGTGCCCAAGGACTTAAAAAAGGCCATGGAGCCAGAGGGGAGGGGCAA 420
QY 4187 CAGACCTTTTCATGGCCAAACCTTTGGGGCCCGTAGTGATCGATTGAACAACCTCGCCAAT 4246
Db 421 CAGACCTTTTCATGGCCAAACCTTTGGGGCCCGTAGTGATCGATTGAACAACCTCGCCAAT 480
QY 4247 CGATACCTCTCTCTTCTTAACCGGACAGGAGGGAACCTCGAGTTTACCACTCCCTATCAGTG 4306
Db 481 CGATCAAGGAGGAGGAGGAGCCAGGACAGGAGGGAAGTGGGAGGGAG----- 525
QY 4307 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCATGATAGAGAAAAGTGAAGT 4366
Db 526 ----- 525
```

```
QY 4367 CGAGTTTACCCTCCCTATCATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTA 4426
Db 526 ----- 525
QY 4427 TCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGT 4486
Db 526 ----- 525
QY 4487 GAAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCAC 4546
Db 526 ----- 525
QY 4547 TCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTCGCTGATCAGCAGAGGACTCCAAAT 4606
Db 526 -----GGTCCAGCAGAGGACTCCAAAT 548
QY 4607 TTAGGCAGCAGGATATGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTGAGAG 4666
Db 549 TTAGGCAGCAGGATATGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTGAGAG 608
QY 4667 ATTTCTCCAAACCAGGATAGAGGAGTTTGGGTGGGGCTCTTCAACCCACACGACCT 4726
Db 609 ATTTCTCCAAACCAGGATAGAGGAGTTTGGGTGGGGCTCTTCAACCCACACGACCT 668
QY 4727 CTCCCACTAGAGAAAAGTGCCTTTCTCGAAGTGGGTTTCAGGCCGCTCAGAGATCT 4786
Db 669 CTCCCACTAGAGAAAAGTGCCTTTCTCGAAGTGGGTTTCAGGCCGCTCAGAGATCT 728
QY 4787 GACAGGGTGGCTTCCACAGCCTGGGAAATCTCAGTGGCAGAGAGTTTCCACAGAAA 4846
Db 729 GACAGGGTGGCTTCCACAGCCTGGGAAATCTCAGTGGCAGAGAGTTTCCACAGAAA 788
QY 4847 CACTGATGCCCTTCCCTTACGCTGTCTTCCATCTTCTTCCCTCGGGATGCTCTCC 4906
Db 789 CACTGATGCCCTTCCCTTACGCTGTCTTCCATCTTCTTCCCTCGGGATGCTCTCC 848
QY 4907 CGTCTGGTTATCTTGGCTCTTGGCTTTCAGCAAGATTGCCCTGTGCTCCACTCCA 4966
Db 849 CGTCTGGTTATCTTGGCTCTTGGCTTTCAGCAAGATTGCCCTGTGCTCCACTCCA 908
QY 4967 TCTTCTCTACTGTCTCCGCTTGCCTTGCCTTCTTGGTGTCTCTTCTTCCACCCA 5026
Db 909 TCTTCTCTACTGTCTCCGCTTGCCTTGCCTTCTTGGTGTCTCTTCTTCCACCCA 968
QY 5027 TTTCTCACTTCACTTTTCTCCCTTCTCATTTGATTCATCTTCTTCTTCTTCTTCT 5086
Db 969 TTTCTCACTTCACTTTTCTCCCTTCTCATTTGATTCATCTTCTTCTTCTTCTTCT 1028
QY 5087 TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5146
Db 1029 TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1088
QY 5147 TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5206
Db 1089 TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
QY 5207 TAAACAATCTCAGTGAGCCACAGCTTCACTGCTGCTGGGTCTCTTACCTTCTCA 5266
Db 1149 TAAACAATCTCAGTGAGCCACAGCTTCACTGCTGCTGGGTCTCTTACCTTCTCA 1208
QY 5267 CCCCCCTGGCTTCTGCTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5326
Db 1209 CCCCCCTGGCTTCTGCTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1268
QY 5327 CTACTCTTCTTCTGCTTGTCTCTCTCTGCTCCAGTCCGACCTGCTGCTGCTGCT 5386
Db 1269 CTACTCTTCTTCTGCTTGTCTCTCTCTGCTCCAGTCCGACCTGCTGCTGCTGCT 1328
QY 5387 CAGCTGTGTGCCACATCTTCTCAGGATCTCTGAAAAGTTTAAACAGGTGAGAAATGTTTCCC 5446
Db 1329 CAGCTGTGTGCCACATCTTCTCAGGATCTCTGAAAAGTTTAAACAGGTGAGAAATGTTTCCC 1388
QY 5447 CTGTAGACAGCAGATCAGATCTTCCGGGAAGTCAGGCTTCCAGCCCTCTCTTCTCTGC 5506
```

```
Db 1389 CTGTAGACAGCAGATCAGATCTTCTCCGGAAGTCAGGCTTCCAGCCCTCTCTTCTCTGC 1448
QY 5507 CCAGCTGTCCCGGCACTCTTTAGCAAACTCAGGACCCCTTACCCCAATAGACCTCTGACA 5566
Db 1449 CCAGCTGTCCCGGCACTCTTTAGCAAACTCAGGACCCCTTACCCCAATAGACCTCTGACA 1508
QY 5567 GAGAACGACGCACTTTTACATGAGTCTGCTGGAGAGCCATAGGCTTACGGTGTAAAGA 5626
Db 1509 GAGAACGACGCACTTTTACATGAGTCTGCTGGAGAGCCATAGGCTTACGGTGTAAAGA 1568
QY 5627 GCAGGGAAGTGTGTGTAGGAAAGTCAGGACTTCCATAGAGCTAGCCACACCCAG 5686
Db 1569 GCAGGGAAGTGTGTGTAGGAAAGTCAGGACTTCCATAGAGCTAGCCACACCCAG 1628
QY 5687 AAATGACAGACAGATCCCTCTCTATCTCCCTCCATAGAGTTTGAGT 5731
Db 1629 AAATGACAGACAGATCCCTCTCTATCTCTCCCTCCATAGAGTTTGAGT 1673
```

RESULT 13
CNS01DTR/c

LOCUS
DEFINITION
Human chromosome 14 DNA sequence BAC C-2201G16 of library CalTech-D
from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION
AL132855
VERSION
AL132855.4 GI:13897280

KEYWORDS

HTG.

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
Sequencing of the human chromosome 14
Unpublished

TITLE

Sequencing of the human chromosome 14

JOURNAL

2 (bases 1 to 157910)

REFERENCE

Genoscope.
Direct Submission
Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

AUTHORS

On Apr 30, 2001 this sequence version replaced gi:12001727.

JOURNAL

- Web : www.genoscope.cns.fr

COMMENT

Center: Genoscope / Centre National de Sequencage

Center code: GS

Web site: <http://www.genoscope.cns.fr/>

Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.

Upstream BAC (overlapping the T7 end) : R-124D2

Downstream BAC (overlapping the SP6 end) : R-66N24 (AC-AL135999)

Assembly program: Phrap; version 2.0

Quality coverage: 6.97x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases

0 : 26

1 - 9 : 8

10 - 19 : 58

20 - 29 : 115

30 - 39 : 574

40 - 49 : 4774

50 - 59 : 7043

60 - 69 : 9171

70 - 79 : 24066

80 - 89 : 56139

90 - 99 : 55936


```
Db 42918 GGCACCTATGACAGCTGAGCGCAGGACAGCA-----ATTCTCTGACAAAAGAAACTGAG 42865
QY 3797 TCGTGCACCTGCAAAAGTGGATCTCTCCCTAGACATCATGACTTTGTCTCTGGGGAGGCA 3856
Db 42864 CCATGGAGATGGACAACAGATCCCTTCCCTGGGCCACCAT-ACTGCAGAGCTTTTAGTCCCTA 42806
QY 3857 GCACCTGTGGAACTTCAGGTCTGAGAGAGTGAAGAGCTCCCTCAGCCTGAGCTATGCAG 3916
Db 42805 GCACCTGGGGCTCCAGTACTACACAGAGAGATGCTCCAGCCTGGGAGCTGTGTGAGGG 42746
QY 3917 ATAGCCAGGGTTGAAAGGGGAGAGGAGAGCCTCGGATGGGAGCTTGTGTGTGGAGGCA 3976
Db 42745 AGGTGAGAATGGGAAGGAGAGCTGGGGAACAGGGGAGGAAAGCCATGTTGGGAGGCG 42686
QY 3977 GGGGACAGATTAATAGCCTGGAAGAGAGTGAACCTTACCAGCTGTGTTCACTCACCT 4036
Db 42685 GAGGACAGGCAATTTGGCTGCGCAGGAGAGGTGACCTCACCATGTTTTCAGTTCCACCT 42626
QY 4037 TCAGATTAATAAATACTGAGGTAAAGGC-----CTGGGTAGGGAGGTGGTGTGAGAGCG 4091
Db 42625 TCGGTTTAAATAAATACTGAGGTAAAGGCCATGGCAGGGTGGGAGAGCGGTGTGAGAGG 42566
QY 4092 TCCTGTCTCTCTCTGTCATGCCCTGAGGCCCTTTGGGAGGAGGAATGTGCCCAAGGACT 4151
Db 42565 TCCTGTCTCTCCACTATCTGCTCATCAGCCCTTTGAGGGAGGAATGTGCCCAAGGACT 42506
QY 4152 AAAAAAGGCCATGAGGAGCAGAGGGCGGAGGCAACAGACCTTTTCATGGGCAAACTTGG 4211
Db 42505 AAAAAAGGCCCTGGAGCCAGAGAGGCTGGGGCAGCAGACCTTTCTTGGGCAAAATCAGG 42446
QY 4212 GGCCC 4216
Db 42445 GGCCC 42441
```

```
RESULT 14
CNS0000B/c
LOCUS
DEFINITION Human chromosome 14 DNA sequence BAC R-124D2 of library RPCI-11
from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL049829
VERSION AL049829.4 GI:8217859
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 196292)
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 196292)
Genoscope.
Direct Submission
Submitted (21-MAY-2001) Genoscope - Centre National de Sequencage :
AP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jun 3, 2000 this sequence version replaced gi:6138746.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-244E17
Downstream BAC (overlapping the SP6 end) : C-2201G16 (AC=AL132855)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.94x in Q20 bases; sum-of-contigs
```

Overall quality chart :

Range : bases

0 : 9

10 : 19

20 : 29

30 : 39

40 : 49

50 : 59

60 : 69

70 : 79

80 : 89

90 : 99

Percentage of bases with a quality value >= 40 : 99 %

Location/Qualifiers

1. 196292

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="14"

/clone="R-124D2"

/clone_lib="RPCI-11"

78259..78398

/note="matching EMBL:AA908790

RHdb:RH102162

dbSTS:STS69699

Identified using the e-PCR software (G. Schuler)"

80433..80566

/note="matching EMBL:AA167748

RHdb:RH98727

dbSTS:STS68485

Identified using the e-PCR software (G. Schuler)"

82164..82299

/note="matching EMBL:H79035

RHdb:RH78773

dbSTS:STS5514

Identified using the e-PCR software (G. Schuler)"

93580..93808

/note="matching EMBL:AA007328

RHdb:RH75103

dbSTS:STS52191

Identified using the e-PCR software (G. Schuler)"

107758..107951

/note="matching EMBL:R94929

RHdb:RH65111

dbSTS:STS45044

Identified using the e-PCR software (G. Schuler)"

107996..108181

/note="matching EMBL:D11677

RHdb:RH47139

dbSTS:STS40201

Identified using the e-PCR software (G. Schuler)"

128387..128525

/note="matching EMBL:H22648

RHdb:RH53520

dbSTS:STS25817

Identified using the e-PCR software (G. Schuler)"

128419..128543

/note="matching EMBL:R87257

RHdb:RH53698

dbSTS:STS18321

Identified using the e-PCR software (G. Schuler)"

138462..138738

/note="matching EMBL:R59134

RHdb:RH53972

dbSTS:STS42930

Identified using the e-PCR software (G. Schuler)"

138509..138658

/note="matching EMBL:M78864

RHdb:RH95543

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 25, 2005, 02:58:32 ; Search time 2127.75 Seconds
(without alignments)
16363.543 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatccgcaagggtcacaca.....ccataagagttgagtcgac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 5735 | 100.0 | 5735 | 17 | US-10-613-728-1 |
| 2 | 5045 | 88.0 | 5443 | 13 | US-10-024-066-5 |
| 3 | 5045 | 88.0 | 5443 | 16 | US-10-332-966-2 |
| 4 | 5045 | 88.0 | 5443 | 19 | US-10-944-375-15 |
| 5 | 1279.8 | 22.3 | 1679 | 17 | US-10-027-655-5 |
| 6 | 407.6 | 7.1 | 5190 | 18 | US-10-798-037-1 |
| 7 | 311.8 | 5.4 | 520 | 9 | US-09-874-389-7 |
| 8 | 311.8 | 5.4 | 520 | 10 | US-09-921-650-7 |
| 9 | 311.8 | 5.4 | 520 | 11 | US-09-241-347-7 |
| 10 | 311.4 | 5.4 | 990 | 15 | US-10-134-643-17 |
| 11 | 311 | 5.4 | 4438 | 16 | US-10-102-143-12 |
| 12 | | | | | Sequence 1, Appli |
| 13 | | | | | Sequence 5, Appli |
| 14 | | | | | Sequence 2, Appli |
| 15 | | | | | Sequence 15, Appli |
| 16 | | | | | Sequence 5, Appli |
| 17 | | | | | Sequence 1, Appli |
| 18 | | | | | Sequence 7, Appli |
| 19 | | | | | Sequence 7, Appli |
| 20 | | | | | Sequence 17, Appli |
| 21 | | | | | Sequence 12, Appli |

| | | | | | | |
|------|-------|-----|------|----|-------------------|--------------------|
| c 12 | 311 | 5.4 | 4479 | 16 | US-10-102-143-10 | Sequence 10, Appli |
| c 13 | 311 | 5.4 | 4556 | 16 | US-10-102-143-8 | Sequence 8, Appli |
| c 14 | 311 | 5.4 | 6345 | 16 | US-10-102-143-6 | Sequence 6, Appli |
| c 15 | 311 | 5.4 | 6423 | 16 | US-10-102-143-4 | Sequence 4, Appli |
| c 16 | 311 | 5.4 | 8287 | 16 | US-10-102-143-14 | Sequence 14, Appli |
| c 17 | 311 | 5.4 | 8364 | 16 | US-10-102-143-16 | Sequence 16, Appli |
| c 18 | 310.4 | 5.4 | 450 | 9 | US-09-281-674-6 | Sequence 6, Appli |
| c 19 | 310.4 | 5.4 | 450 | 9 | US-09-777-317-6 | Sequence 6, Appli |
| c 20 | 310.4 | 5.4 | 450 | 9 | US-09-892-227-6 | Sequence 6, Appli |
| c 21 | 310.4 | 5.4 | 450 | 9 | US-09-874-389-9 | Sequence 9, Appli |
| c 22 | 310.4 | 5.4 | 450 | 10 | US-09-921-650-9 | Sequence 9, Appli |
| c 23 | 310.4 | 5.4 | 450 | 11 | US-09-241-347-9 | Sequence 9, Appli |
| c 24 | 309.4 | 5.4 | 447 | 19 | US-10-169-050-3 | Sequence 3, Appli |
| c 25 | 309.4 | 5.4 | 450 | 9 | US-09-281-674-5 | Sequence 5, Appli |
| c 26 | 309.4 | 5.4 | 450 | 9 | US-09-777-317-5 | Sequence 5, Appli |
| c 27 | 309.4 | 5.4 | 450 | 9 | US-09-892-227-5 | Sequence 5, Appli |
| c 28 | 309.4 | 5.4 | 450 | 9 | US-09-874-389-8 | Sequence 8, Appli |
| c 29 | 309.4 | 5.4 | 450 | 10 | US-09-921-650-8 | Sequence 8, Appli |
| c 30 | 309.4 | 5.4 | 450 | 11 | US-09-241-347-8 | Sequence 8, Appli |
| c 31 | 309.4 | 5.4 | 456 | 9 | US-09-900-530A-28 | Sequence 28, Appli |
| c 32 | 309.4 | 5.4 | 470 | 16 | US-10-375-884-2 | Sequence 2, Appli |
| c 33 | 309.4 | 5.4 | 500 | 17 | US-10-148-521-21 | Sequence 21, Appli |
| c 34 | 309.4 | 5.4 | 2061 | 18 | US-10-789-303-1 | Sequence 1, Appli |
| c 35 | 309.4 | 5.4 | 2792 | 16 | US-10-375-884-8 | Sequence 8, Appli |
| c 36 | 309.4 | 5.4 | 3621 | 10 | US-09-957-458B-7 | Sequence 7, Appli |
| c 37 | 309.4 | 5.4 | 3752 | 10 | US-09-957-458B-8 | Sequence 8, Appli |
| c 38 | 309.4 | 5.4 | 3871 | 19 | US-10-169-050-51 | Sequence 51, Appli |
| c 39 | 309.4 | 5.4 | 3926 | 19 | US-10-169-050-59 | Sequence 59, Appli |
| c 40 | 309.4 | 5.4 | 3930 | 19 | US-10-169-050-36 | Sequence 36, Appli |
| c 41 | 309.4 | 5.4 | 4224 | 10 | US-09-957-458B-10 | Sequence 10, Appli |
| c 42 | 309.4 | 5.4 | 4382 | 10 | US-09-957-458B-9 | Sequence 9, Appli |
| c 43 | 309.4 | 5.4 | 4386 | 19 | US-10-169-050-57 | Sequence 57, Appli |
| c 44 | 309.4 | 5.4 | 4824 | 19 | US-10-169-050-52 | Sequence 52, Appli |
| c 45 | 309.4 | 5.4 | 4963 | 9 | US-09-281-674-9 | Sequence 9, Appli |

ALIGNMENTS

RESULT 1
US-10-613-728-1
; Sequence 1, Application US/10613728
; Publication No. US20040010813A1
; GENERAL INFORMATION:
; APPLICANT: Cincinnati Children's Hospital Medical Center
; APPLICANT: Robbins, Jeffrey
; TITLE OF INVENTION: A ROBUST, INDUCIBLE CARDIAC PREFERRED
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR TRANSGENESIS
; FILE REFERENCE: CHM02 GN053
; CURRENT APPLICATION NUMBER: US/10/613,728
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: 60/393,525
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5735
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Inducible Cardiac preferred promoter
US-10-613-728-1

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 100.0% | Score 5735; | DB 17; | Length 5735; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 5735; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | GGATCTCGAAGGTTCACACAGAGGTCTCCACCACAGTGCCTAGTCTCAATTCAGT | 60 | |
| Db | 1 | GGATCTCGAAGGTTCACACAGAGGTCTCCACCACAGTGCCTAGTCTCAATTCAGT | 60 | |
| QY | 61 | TTCCATGCTTGTCTCAATGCTGGCTCCACAGAGTAAATTTGGACTTTGTTTAT | 120 | |

Db 61 TTCCATGCTGTTTCTCAATGCTGGCTCCCCAGAGCTAATTTGGACTTTGTTTAT 120
QY 121 TTCAAAAGGCGCTGAATGAGGAGTAGATCTTGCTGTACCCAGCTCTAAGGGTGCCCGTA 180
Db 121 TTCAAAAGGCGCTGAATGAGGAGTAGATCTTGCTGTACCCAGCTCTAAGGGTGCCCGTA 180
QY 181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAAATAAGCAATTT 240
Db 181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAAATAAGCAATTT 240
QY 241 TCCTTAAAGCMAATCTGCGCTCTAGACTCTTCTCTGTGACCTCGGTCCCTGGGCTCT 300
Db 241 TCCTTAAAGCMAATCTGCGCTCTAGACTCTTCTCTGTGACCTCGGTCCCTGGGCTCT 300
QY 301 AGGGTGGGAGGTGGGCTTGGGAAGAAGGTGGGAAGTGGCAAGCCGATCCCTAG 360
Db 301 AGGGTGGGAGGTGGGCTTGGGAAGAAGGTGGGAAGTGGCAAGCCGATCCCTAG 360
QY 361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420
Db 361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420
QY 421 AAACATAGCAAGATGATACCTCTTTGTGTATCTTCCGAGCCAGTACCTGTGAGGTT 480
Db 421 AAACATAGCAAGATGATACCTCTTTGTGTATCTTCCGAGCCAGTACCTGTGAGGTT 480
QY 481 GAAACAGGATTTAGAGAGCCCTCGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
Db 481 GAAACAGGATTTAGAGAGCCCTCGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
QY 541 AGCACTAGGTGCCACTGCTAGTATCTTACGCTGATATATGACAGAGCTGGGCCAC 600
Db 541 AGCACTAGGTGCCACTGCTAGTATCTTACGCTGATATATGACAGAGCTGGGCCAC 600
QY 601 AGAAGTCTGGGGTGTAGAACTGACAGTGAATTTTTCAGTGGCAAGGATGACCCCC 660
Db 601 AGAAGTCTGGGGTGTAGAACTGACAGTGAATTTTTCAGTGGCAAGGATGACCCCC 660
QY 661 TCAGCAGATGTAGTATGTCCTTTAGATCCATCCAGGAGGCTCTTACAGGACATG 720
Db 661 TCAGCAGATGTAGTATGTCCTTTAGATCCATCCAGGAGGCTCTTACAGGACATG 720
QY 721 GGATGAGAGATGTAGTATGTCCTTTAGATCCATCCAGGAGGCTCTTACAGGACATG 780
Db 721 GGATGAGAGATGTAGTATGTCCTTTAGATCCATCCAGGAGGCTCTTACAGGACATG 780
QY 781 TTCCACTTAGCCAGAGAGCAGTAACTTACCTATCTTCTCCCTCCCATCTCCAG 840
Db 781 TTCCACTTAGCCAGAGAGCAGTAACTTACCTATCTTCTCCCTCCCATCTCCAG 840
QY 841 GACACACCCCTGGTCTGACGATTTCAATTTCTTCTTACGCTCCCTCTGTGACTTCCAT 900
Db 841 GACACACCCCTGGTCTGACGATTTCAATTTCTTCTTACGCTCCCTCTGTGACTTCCAT 900
QY 901 TTGCAAGGCTTTTGACCTCTGACGCTGTGGAAGATAGAGTTTGGCCCTAGGTGGCAA 960
Db 901 TTGCAAGGCTTTTGACCTCTGACGCTGTGGAAGATAGAGTTTGGCCCTAGGTGGCAA 960
QY 961 GCCATCTCAAGAGAAAGCAGACAAACAGGGGGACAGATTTTGGAAAGGATCAGGAATAA 1020
Db 961 GCCATCTCAAGAGAAAGCAGACAAACAGGGGGACAGATTTTGGAAAGGATCAGGAATAA 1020
QY 1021 TCACCTGGGGGCTGGGGGTAGAAAAGAGTGTAGTCCGCTCCAGCTAAGCCAGC 1080
Db 1021 TCACCTGGGGGCTGGGGGTAGAAAAGAGTGTAGTCCGCTCCAGCTAAGCCAGC 1080
QY 1081 TAGTCCCCAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTC 1140
Db 1081 TAGTCCCCAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTC 1140
QY 1141 TGAAGAACAATGGGATTTGGAAGACATCTTTTGTAGTCTCCCTCAACCCCACTACAGAC 1200
Db 1141 TGAAGAACAATGGGATTTGGAAGACATCTTTTGTAGTCTCCCTCAACCCCACTACAGAC 1200

QY 1201 AACTCGTGTGTGGCCAGACTCCTGTTCAACAGGCCCTCTGTGTTCTGACCACTGAGCTAG 1260
Db 1201 AACTCGTGTGTGGCCAGACTCCTGTTCAACAGGCCCTCTGTGTTCTGACCACTGAGCTAG 1260
QY 1261 GCAACCCAGAGCATGGGCCCTGTGTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG 1320
Db 1261 GCAACCCAGAGCATGGGCCCTGTGTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG 1320
QY 1321 CAGGGGAGGAGAAACAGAGAAACGAAATTAAGAAAGGAAAGGCAAGTCAATCAGA 1380
Db 1321 CAGGGGAGGAGAAACAGAGAAACGAAATTAAGAAAGGAAAGGCAAGTCAATCAGA 1380
QY 1381 TGCAGTCAAGAGAGATGGGAAGCCAAACACAGCTTTGAGCAGAGAAACAGAAAGGAG 1440
Db 1381 TGCAGTCAAGAGAGATGGGAAGCCAAACACAGCTTTGAGCAGAGAAACAGAAAGGAG 1440
QY 1441 AGATTTCTGGGCATTAAGGAGGCCACAGAAAGAGAGCCAGGCCCCCAAGTCTCCTTTT 1500
Db 1441 AGATTTCTGGGCATTAAGGAGGCCACAGAAAGAGAGCCAGGCCCCCAAGTCTCCTTTT 1500
QY 1501 ATACCTCATCCCGTCTCCCAATTAAGCCCACTCTTTCTTCTAGATCAGACCTGAGCTGC 1560
Db 1501 ATACCTCATCCCGTCTCCCAATTAAGCCCACTCTTTCTTCTAGATCAGACCTGAGCTGC 1560
QY 1561 AGCGAAGAGAGCCCTAGGGAGGATCACACTGATGAAGAGATGTGTGGAAGTCCAGG 1620
Db 1561 AGCGAAGAGAGCCCTAGGGAGGATCACACTGATGAAGAGATGTGTGGAAGTCCAGG 1620
QY 1621 GCAACCTTAAGAGCCAGAGCCTTAAAGAGCAAGAGATTAAGGTGCTTCAAGGTGGCCAGG 1680
Db 1621 GCAACCTTAAGAGCCAGAGCCTTAAAGAGCAAGAGATTAAGGTGCTTCAAGGTGGCCAGG 1680
QY 1681 CTGTGCAACAGAGGCTGAGGACTGTGTGTAGAGCTCAAGATTAAGATGATGCTCAGA 1740
Db 1681 CTGTGCAACAGAGGCTGAGGACTGTGTGTAGAGCTCAAGATTAAGATGATGCTCAGA 1740
QY 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAGAGTGAAGAGCCTGGAAACAG 1800
Db 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAGAGTGAAGAGCCTGGAAACAG 1800
QY 1801 AGAATCTGGAGCGCTGGAAACGATACCATTAAGGGAGAAACCCAGGCTACCTTTAGATG 1860
Db 1801 AGAATCTGGAGCGCTGGAAACGATACCATTAAGGGAGAAACCCAGGCTACCTTTAGATG 1860
QY 1861 TAAATCATGAAGAACAGGGAGAGGAGCTGGAGAGATGAAGGACCCCGGGGCAAGA 1920
Db 1861 TAAATCATGAAGAACAGGGAGAGGAGCTGGAGAGATGAAGGACCCCGGGGCAAGA 1920
QY 1921 CATGGAAGCAGGACAAAGCAGAGGCTAGGCTCCGTGAAATCAGCCTCTGAAGGCAGAG 1980
Db 1921 CATGGAAGCAGGACAAAGCAGAGGCTAGGCTCCGTGAAATCAGCCTCTGAAGGCAGAG 1980
QY 1981 CCCTGCTATGAGCACCAGAAACAGAGAGGCTAGGCTTAATGTCGAGACAGGGAAACAGAG 2040
Db 1981 CCCTGCTATGAGCACCAGAAACAGAGAGGCTAGGCTTAATGTCGAGACAGGGAAACAGAG 2040
QY 2041 GTAGACACAGGAAACAGACAGAGACGGGGGAGCCAGGTAAACAAAGAAATGGTCTCTCAC 2100
Db 2041 GTAGACACAGGAAACAGACAGAGACGGGGGAGCCAGGTAAACAAAGAAATGGTCTCTCAC 2100
QY 2101 CTGTGGCCAGAGCCTCATCTGTCCACATCTAGAAATGTTTATCAGACTGCGAGGC 2160
Db 2101 CTGTGGCCAGAGCCTCATCTGTCCACATCTAGAAATGTTTATCAGACTGCGAGGC 2160
QY 2161 TGGCTTGGGAGCAGCTGGAAAGATGTGTAGAGCCAGGGGAGACAGGGGGCTTAGGA 2220
Db 2161 TGGCTTGGGAGCAGCTGGAAAGATGTGTAGAGCCAGGGGAGACAGGGGGCTTAGGA 2220
QY 2221 AAGGAAGAGAGGGGCAACAGGGCCACAAAGAGGGGAGAGCCCGGAACTGAGTTAACTC 2280
Db 2221 AAGGAAGAGAGGGGCAACAGGGCCACAAAGAGGGGAGAGCCCGGAACTGAGTTAACTC 2280

QY 2281 CTTCTTGTGTGATCTTCCATAGGAGGAGTGGGAACCTGTGACCAACCATCCCCATGA 2340
Db |||||
QY 2281 CTTCTTGTGTGATCTTCCATAGGAGGAGTGGGAACCTGTGACCAACCATCCCCATGA 2340
Db |||||
QY 2341 GCCCCACTACCCATACCAAGTTTGGCTGAGTGGCACTCTAGGTTCCCTGAGGACAGAG 2400
Db |||||
QY 2341 GCCCCACTACCCATACCAAGTTTGGCTGAGTGGCACTCTAGGTTCCCTGAGGACAGAG 2400
Db |||||
QY 2401 CTTGGCTTTTGTCTCTGGACCTGACCAAGCTGACCAAGTCTCAGTACCTTTGAAT 2460
Db |||||
QY 2401 CTTGGCTTTTGTCTCTGGACCTGACCAAGCTGACCAAGTCTCAGTACCTTTGAAT 2460
Db |||||
QY 2461 GCCCTCAAGAGCTTGAGAACCAAGGAGTGCATATTTAGGCCCATGAGGCTTATGCC 2520
Db |||||
QY 2461 GCCCTCAAGAGCTTGAGAACCAAGGAGTGCATATTTAGGCCCATGAGGCTTATGCC 2520
Db |||||
QY 2521 TGCACACAGGAGCTCAAGTACCTCCAGGGACACAGCTGCACAGAGTGGCTTTATCC 2580
Db |||||
QY 2521 TGCACACAGGAGCTCAAGTACCTCCAGGGACACAGCTGCACAGAGTGGCTTTATCC 2580
Db |||||
QY 2581 CCAAGAGCAACCAATTTGGCATAGTGTGCTGCAATGGGATGCAAGTTGNAATCAGGTC 2640
Db |||||
QY 2581 CCAAGAGCAACCAATTTGGCATAGTGTGCTGCAATGGGATGCAAGTTGNAATCAGGTC 2640
Db |||||
QY 2641 CCTTCAAGATACTGCATGCAAGACCTAAGACCCCTGGAGAGAGGGTATGCTCTGCC 2700
Db |||||
QY 2641 CCTTCAAGATACTGCATGCAAGACCTAAGACCCCTGGAGAGAGGGTATGCTCTGCC 2700
Db |||||
QY 2701 CCACCCACATAAGGGAGTGAATCTCTAGGGGCTGGGACCTTGGGGAGACACAC 2760
Db |||||
QY 2701 CCACCCACATAAGGGAGTGAATCTCTAGGGGCTGGGACCTTGGGGAGACACAC 2760
Db |||||
QY 2761 ATTACTGAGAGTGTGAGCCCAAGAACTGACCGCTGTGCTGCTGCCACCTCCACAC 2820
Db |||||
QY 2761 ATTACTGAGAGTGTGAGCCCAAGAACTGACCGCTGTGCTGCTGCCACCTCCACAC 2820
Db |||||
QY 2821 TCTAGAGCTATATTGAGAGGTGACATAGATAGAGTGGGAGCTGTAGCAGGAGAGTGT 2880
Db |||||
QY 2821 TCTAGAGCTATATTGAGAGGTGACATAGATAGAGTGGGAGCTGTAGCAGGAGAGTGT 2880
Db |||||
QY 2881 TCCTGGGTGTGAGGGTGTAGGGGAAGCCAGACAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Db |||||
QY 2881 TCCTGGGTGTGAGGGTGTAGGGGAAGCCAGACAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Db |||||
QY 2941 CACATGTCTACTTGTATTATACAGGAGTACCTGCTAAGACCCCAACATCTACGACCTC 3000
Db |||||
QY 2941 CACATGTCTACTTGTATTATACAGGAGTACCTGCTAAGACCCCAACATCTACGACCTC 3000
Db |||||
QY 3001 TGAAGAAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGGCTTTGGGTGTGATGGTG 3060
Db |||||
QY 3001 TGAAGAAGACAGAGCCCTGGAGGACAGGGGTTGTCTCTGAGGCTTTGGGTGTGATGGTG 3060
Db |||||
QY 3061 CCACAAAGGAGGAGTGTGAGTATAGGCCCCCAGGAGCGTTAGAGAGGGGACACTTG 3120
Db |||||
QY 3061 CCACAAAGGAGGAGTGTGAGTATAGGCCCCCAGGAGCGTTAGAGAGGGGACACTTG 3120
Db |||||
QY 3121 GGAAGGGTGTGAGTGTGAGAGCCCTATTCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
Db |||||
QY 3121 GGAAGGGTGTGAGTGTGAGAGCCCTATTCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
Db |||||
QY 3181 TAAATCTCTGGGCTGCCAGGCAATTCAGAGGACCTGCACTCTCTGCGAGGCTGGGGA 3240
Db |||||
QY 3181 TAAATCTCTGGGCTGCCAGGCAATTCAGAGGACCTGCACTCTCTGCGAGGCTGGGGA 3240
Db |||||
QY 3241 GGCGGAAGGAGCAACCCCTTATACCTTTCTCCCTCAGCCCCAGGATTAACACT 3300
Db |||||
QY 3241 GGCGGAAGGAGCAACCCCTTATACCTTTCTCCCTCAGCCCCAGGATTAACACT 3300
Db |||||
QY 3301 CTGGCTTTCCCTTTCCACTTCCATCAGAGTGGAGGGTTGACAGAGGGGTAAGAAA 3360
Db |||||
QY 3301 CTGGCTTTCCCTTTCCACTTCCATCAGAGTGGAGGGTTGACAGAGGGGTAAGAAA 3360
Db |||||
QY 3361 CCTACATGTCCAAACATCATGGTGCACCATATATGATCAGTATGTGTAGAGGCAAGAAA 3420

Db |||||
QY 3361 CCTACATGTCCAAACATCATGGTGCACCATATATGATCAGTATGTGTAGAGGCAAGAAA 3420
Db |||||
QY 3421 GGAATCTGCAAGGCTTAACTGGTTAAATGTGTAAAGTCTGTGTGATGTGTGTGTCTG 3480
Db |||||
QY 3421 GGAATCTGCAAGGCTTAACTGGTTAAATGTGTAAAGTCTGTGTGATGTGTGTGTCTG 3480
Db |||||
QY 3481 ACTGAAAAACGGCATGGCTGTGACGTCTCAGTTCTGTGCGTGAAGTTTACCAGACTGCA 3540
Db |||||
QY 3481 ACTGAAAAACGGCATGGCTGTGACGTCTCAGTTCTGTGCGTGAAGTTTACCAGACTGCA 3540
Db |||||
QY 3541 GGTTTGTGTGTAATTTGCCAAGGCAAAAGTGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
Db |||||
QY 3541 GGTTTGTGTGTAATTTGCCAAGGCAAAAGTGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
Db |||||
QY 3601 TGGATGATGGCTGCACTCTCAAGGACCATGGAANAATAGAAATGGAACCTCTATATGTCT 3660
Db |||||
QY 3601 TGGATGATGGCTGCACTCTCAAGGACCATGGAANAATAGAAATGGAACCTCTATATGTCT 3660
Db |||||
QY 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
Db |||||
QY 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
Db |||||
QY 3721 TACAGACAGTATCTGTACAGAGTAAAGGAGAGAGGGGGTGTAGAAATCTCTTACTA 3780
Db |||||
QY 3721 TACAGACAGTATCTGTACAGAGTAAAGGAGAGAGGGGGTGTAGAAATCTCTTACTA 3780
Db |||||
QY 3781 TCAAAGGGAATCTGAGTGTGACCTGCAAAAGTGAATCTCTCTAGACATCATGACTT 3840
Db |||||
QY 3781 TCAAAGGGAATCTGAGTGTGACCTGCAAAAGTGAATCTCTCTAGACATCATGACTT 3840
Db |||||
QY 3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGGCTGAGAGAGTGGAGGCTCCCTCTCA 3900
Db |||||
QY 3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGGCTGAGAGAGTGGAGGCTCCCTCTCA 3900
Db |||||
QY 3901 GCCTGAAGCTATGAGATAGCAGGCTTGAAGGGGGAAGGAGAGCTTGGGATGGGAGC 3960
Db |||||
QY 3901 GCCTGAAGCTATGAGATAGCAGGCTTGAAGGGGGAAGGAGAGCTTGGGATGGGAGC 3960
Db |||||
QY 3961 TTGTGTGTGGAGGAGGAGGAGCAGATATTAAGCTTGAAGAGAGAGGAGTGAACCTTACCAG 4020
Db |||||
QY 3961 TTGTGTGTGGAGGAGGAGGAGCAGATATTAAGCTTGAAGAGAGAGGAGTGAACCTTACCAG 4020
Db |||||
QY 4021 TTGTTCACCTCACCTTTCAGATTAATAAATACTGAGTAAAGGCTGGTGGGAGGAGTG 4080
Db |||||
QY 4021 TTGTTCACCTCACCTTTCAGATTAATAAATACTGAGTAAAGGCTGGTGGGAGGAGTG 4080
Db |||||
QY 4081 GTGTGAGAGCTCTCTCTCTCTGATGCTTGGAGGCTTGGGAGGAGGAGTGT 4140
Db |||||
QY 4081 GTGTGAGAGCTCTCTCTCTCTGATGCTTGGAGGCTTGGGAGGAGGAGTGT 4140
Db |||||
QY 4141 GCGGAGGAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
Db |||||
QY 4141 GCGGAGGAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
Db |||||
QY 4201 GCAAACTTGGGGGCGGTGAGTGTGATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 4260
Db |||||
QY 4201 GCAAACTTGGGGGCGGTGAGTGTGATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 4260
Db |||||
QY 4261 TTCTAAACGAGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4320
Db |||||
QY 4261 TTCTAAACGAGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4320
Db |||||
QY 4321 AAGTCGAGTTTACCACTCCCTTATCAGTATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380
Db |||||
QY 4321 AAGTCGAGTTTACCACTCCCTTATCAGTATAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380
Db |||||
QY 4381 CCTATCAGTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4440
Db |||||
QY 4381 CCTATCAGTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4440
Db |||||
QY 4441 AAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4500

```
Db 4441 AAGTGAAGTCGAGTTTACCCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4500
Qy 4501 CCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGAT 4560
Db 4501 CCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCCTCCCTATCAGTGAT 4560
Qy 4561 AGAGAAAAGTGAAAGTCGAGTCGGTACCAGCAGAGGACTCCAAATTTAGGCAGCAGGCA 4620
Db 4561 AGAGAAAAGTGAAAGTCGAGTCGGTACCAGCAGAGGACTCCAAATTTAGGCAGCAGGCA 4620
Qy 4621 TATGGGATGGGATATAAAGGGGCTGGAGCAGTGTGAGAGATTTCTCCAAACCCA 4680
Db 4621 TATGGGATGGGATATAAAGGGGCTGGAGCAGTGTGAGAGATTTCTCCAAACCCA 4680
Qy 4681 GGTAAAGGGAGTTTCGGGTGGGGGCTCTTCAACCCACACAGACCTCTCCCCACCTAGAA 4740
Db 4681 GGTAAAGGGAGTTTCGGGTGGGGGCTCTTCAACCCACACAGACCTCTCCCCACCTAGAA 4740
Qy 4741 GGAAACTGCCCTTCTCGGAAGTGCGGTTTCAGGCCGCTCAGAGATCTGACAGGGTGGCCCT 4800
Db 4741 GGAAACTGCCCTTCTCGGAAGTGCGGTTTCAGGCCGCTCAGAGATCTGACAGGGTGGCCCT 4800
Qy 4801 CCACAGCTGGGAAGTTCTAGTGGCAGAGGTTTCCACAGAAACACTGAGTGCCTT 4860
Db 4801 CCACAGCTGGGAAGTTCTAGTGGCAGAGGTTTCCACAGAAACACTGAGTGCCTT 4860
Qy 4861 TCCCTTACGCTGCTCTTCCATCTTCCCTGCGGATGCTCTCCCGCTCTTGGTTTATC 4920
Db 4861 TCCCTTACGCTGCTCTTCCATCTTCCCTGCGGATGCTCTCCCGCTCTTGGTTTATC 4920
Qy 4921 TTGGCTCTTCTGCTTTCAGCAAGATTTGGCCCTGTGCTGCCACTCCATCTTCTCTACTGT 4980
Db 4921 TTGGCTCTTCTGCTTTCAGCAAGATTTGGCCCTGTGCTGCCACTCCATCTTCTCTACTGT 4980
Qy 4981 CTCGGTGCCTTGCCTTCTTGGTGTCTTCCCTTCCAGCCATTTCTCACTTACC 5040
Db 4981 CTCGGTGCCTTGCCTTCTTGGTGTCTTCCCTTCCAGCCATTTCTCACTTACC 5040
Qy 5041 TTTTCTCCCTCTCATTTGTAATCATCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTC 5100
Db 5041 TTTTCTCCCTCTCATTTGTAATCATCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTC 5100
Qy 5101 CTTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5160
Db 5101 CTTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5160
Qy 5161 CTGTGTACAGTGTGAGAAACACACCTGGGGTTCACACCTTATATGTAACAAATCTTCCA 5220
Db 5161 CTGTGTACAGTGTGAGAAACACACCTGGGGTTCACACCTTATATGTAACAAATCTTCCA 5220
Qy 5221 GTGAGCCACAGCTTCAGTGTGCTGGGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5280
Db 5221 GTGAGCCACAGCTTCAGTGTGCTGGGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5280
Qy 5281 CTGTTTCCATCTCTGTTTCTGAGATCTCTGATTTGTTCTTCCAGGCTCTGTCTTCTTCT 5340
Db 5281 CTGTTTCCATCTCTGTTTCTGAGATTTGTTCTTCCAGGCTCTGTCTTCTTCTTCTTCT 5340
Qy 5341 GCCTGTTTCTTCTTCTTCTGTCAGCTCGGCACTGTGGTGTCTGTTTCCAGCTGTGTGTCAC 5400
Db 5341 GCCTGTTTCTTCTTCTTCTGTCAGCTCGGCACTGTGGTGTCTGTTTCCAGCTGTGTGTCAC 5400
Qy 5401 ATTCTTCAGGATCTCTGAAAGTTTAAACAGGTGAGATGTTTCCCTGTAGACAGCAGA 5460
Db 5401 ATTCTTCAGGATCTCTGAAAGTTTAAACAGGTGAGATGTTTCCCTGTAGACAGCAGA 5460
Qy 5461 TCACGATTTCTCCCGAAGTCAGGCTTCCAGGCTCTCTTCTTCTTCTGCCCAGCTGCCGCA 5520
Db 5461 TCACGATTTCTCCCGAAGTCAGGCTTCCAGGCTCTCTTCTTCTTCTGCCCAGCTGCCGCA 5520
Qy 5521 CTCTTAGCAAACTCAGGCACTTACCCCAATAGACCTCTGACAGAGAAAGCAGGCACT 5580
Db 5521 CTCTTAGCAAACTCAGGCACTTACCCCAATAGACCTCTGACAGAGAAAGCAGGCACT 5580
```

```
Qy 5581 TTACATCGAGTCTCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAAGAGCAGGGAAGTGGT 5640
Db 5581 TTACATCGAGTCTCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAAGAGCAGGGAAGTGGT 5640
Qy 5641 GGTGTAGGAAAGTCAGGACTTTCACATAGAGAGCCCTAGCCCAACCCAGAAATACACAGACA 5700
Db 5641 GGTGTAGGAAAGTCAGGACTTTCACATAGAGAGCCCTAGCCCAACCCAGAAATACACAGACA 5700
Qy 5701 TCCCTCTCTATCTCCCTCCCATAGAGTTTGAATCGAC 5735
Db 5701 TCCCTCTCTATCTCCCTCCCATAGAGTTTGAATCGAC 5735
```

RESULT 2

US-10-024-066-5

; Sequence 5, Application US/10024066

; Publication No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pagumathi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024,066

; CURRENT FILING DATE: 2001-12-18

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 5443

; TYPE: DNA

; ORGANISM: Mus musculus

US-10-024-066-5

Query Match 88.0%; Score 5045; DB 13; Length 5443;

Best Local Similarity 93.9%; Pred. No. 0;

Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

```
Qy 1 GGATCTCTGCAAGGTCACACAAGGCTCTCCACCCACCAAGGTCCTCAATTTCAGT 60
Db 1 GGATCTCTGCAAGGTCACACAAGGCTCTCCACCCACCAAGGTCCTCAATTTCAGT 60
Qy 61 TTCCATGCTTGTCTTCTCACAATGCTGGCTCTCCAGAGCTAATTTGGACTTTGTTTAT 120
Db 61 TTCCATGCTTGTCTTCTCACAATGCTGGCTCTCCAGAGCTAATTTGGACTTTGTTTAT 120
Qy 121 TTCAAAAAGGGCTCAATGAGGAGTAGATCTTGTCTACCCAGCTCTAAGGTCGCCGTGA 180
Db 121 TTCAAAAAGGGCTCAATGAGGAGTAGATCTTGTCTACCCAGCTCTAAGGTCGCCGTGA 180
Qy 181 AGCCCTCAGACCTCGAGGCTTTTGCAACAGCCCTTTAGTGTGAAGCAGAAATTAAGCAATTT 240
Db 181 AGCCCTCAGACCTCGAGGCTTTTGCAACAGCCCTTTAGTGTGAAGCAGAAATTAAGCAATTT 240
Qy 241 TCCTTAAAGCCAAATCTCTGCTCTAGACTCTTCTCTGACCTCGTCTCCCTGGGCTCT 300
Db 241 TCCTTAAAGCCAAATCTCTGCTCTAGACTCTTCTCTGACCTCGTCTCCCTGGGCTCT 300
Qy 301 AGGGTGGGAGGTCGGGCTTGGAAAGAGAGAGTGGGGAAGTGGAAGCCGATCCCTAG 360
Db 301 AGGGTGGGAGGTCGGGCTTGGAAAGAGAGAGTGGGGAAGTGGAAGCCGATCCCTAG 360
Qy 361 GGCCCTGTGAAGTTCCGAGGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420
Db 361 GGCCCTGTGAAGTTCCGAGGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCCAGCC 420
Qy 421 AAACATAGCAAGAGTAGATACCTCTTTTGTGACTTCCCGCAGGCCAGTACCTGTGAGTT 480
Db 421 AAACATAGCAAGAGTAGATACCTCTTTTGTGACTTCCCGCAGGCCAGTACCTGTGAGTT 480
```

QY 481 GAAACAGGATTTAGAGAGCCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
DB |||||
QY 481 GAAACAGGATTTAGAGAGCCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
DB |||||
QY 541 AGCACCTAGGTGCCACTCTAGTTAGTATCTCTACGCTGATATATGACAGAGCTGGGCCAC 600
DB |||||
QY 541 AGCACCTAGGTGCCACTCTAGTTAGTATCTCTACGCTGATATATGACAGAGCTGGGCCAC 600
DB |||||
QY 601 AGAAGTCTGGGGGTAGAGAACTGACAGTGAATCTTTTCAGTCGGCAAGGTATGACCCCC 660
DB |||||
QY 601 AGAAGTCTGGGGGTAGAGAACTGACAGTGAATCTTTTCAGTCGGCAAGGTATGACCCCC 660
DB |||||
QY 661 TCAGCAGATGTAGTAATGTCTCCCTTAGATCCATCCAGGAGGTCTCTAAGAGGACATG 720
DB |||||
QY 661 TCAGCAGATGTAGTAATGTCTCCCTTAGATCCATCCAGGAGGTCTCTAAGAGGACATG 720
DB |||||
QY 721 GGATGAGAGATGTAGTCAATGTGGCATTTCCAAACACAGCTATCCACAGTGTCCCTTGGCCC 780
DB |||||
QY 781 TTCACCTTAGCCAGGAGACAGTAACCTTAGCTATCTTTCTCTCCCAATCTCTCCAG 840
DB |||||
QY 781 TTCACCTTAGCCAGGAGACAGTAACCTTAGCTATCTTTCTCTCCCAATCTCTCCAG 840
DB |||||
QY 841 GACACACCCCTGGTCTGAGTATTTCTTTCTTCCCTTCAAGTCCCTCTGTGACTTCCAT 900
DB |||||
QY 841 GACACACCCCTGGTCTGAGTATTTCTTTCTTCCCTTCAAGTCCCTCTGTGACTTCCAT 900
DB |||||
QY 901 TTGCAAGGCTTTTGACCTCTGAGCTGCTGGAAGATAGAGTTTGGCCCTTAGGTGGCAA 960
DB |||||
QY 901 TTGCAAGGCTTTTGACCTCTGAGCTGCTGGAAGATAGAGTTTGGCCCTTAGGTGGCAA 960
DB |||||
QY 961 GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGAGTTTGGAAAGATAGGAACTAAA 1020
DB |||||
QY 961 GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGAGTTTGGAAAGATAGGAACTAAA 1020
DB |||||
QY 1021 TCACCTGGGGGCTGGGGGTAGAAAAGAGTGAAGTCCGCTCCAGCTTAAGCCAAAGC 1080
DB |||||
QY 1021 TCACCTGGGGGCTGGGGGTAGAAAAGAGTGAAGTCCGCTCCAGCTTAAGCCAAAGC 1080
DB |||||
QY 1081 TAGTCCCCGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
DB |||||
QY 1081 TAGTCCCCGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
DB |||||
QY 1141 TGAAGACAATGGGATGGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
DB |||||
QY 1141 TGAAGACAATGGGATGGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
DB |||||
QY 1201 ACACCTGTTGTGGCCAGACTCTGTTCAACAGCCCTCTGTTCTGACCACTGAGCTAG 1260
DB |||||
QY 1201 ACACCTGTTGTGGCCAGACTCTGTTCAACAGCCCTCTGTTCTGACCACTGAGCTAG 1260
DB |||||
QY 1261 GCAACCCAGACATGGGCCCTGCTGTGAGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
DB |||||
QY 1261 GCAACCCAGACATGGGCCCTGCTGTGAGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
DB |||||
QY 1321 CAGGGGAGGAGAAACAGAGAAACGAAATGAAGAGGAAAGGAGGAGGCAATCAATCAGA 1380
DB |||||
QY 1321 CAGGGGAGGAGAAACAGAGAAACGAAATGAAGAGGAAAGGAGGAGGCAATCAATCAGA 1380
DB |||||
QY 1381 TGCAGTCAAGAGAGATGGGAAGCCAAACACAGCTTGAAGAGGAAACAGAAAGGGAG 1440
DB |||||
QY 1381 TGCAGTCAAGAGAGATGGGAAGCCAAACACAGCTTGAAGAGGAAACAGAAAGGGAG 1440
DB |||||
QY 1441 AGATTCTGGGATAGGAGGCCACAGAAAGAGAGGCCAGGCCCCCAAGTCTCTCTTTT 1500
DB |||||
QY 1441 AGATTCTGGGATAGGAGGCCACAGAAAGAGAGGCCAGGCCCCCAAGTCTCTCTTTT 1500
DB |||||
QY 1501 ATACCTCATCCCGTCTCCCAATTAAGCCCACTCTTTCTCTAGATCAGACTGAGCTGC 1560
DB |||||
QY 1501 ATACCTCATCCCGTCTCCCAATTAAGCCCACTCTTTCTCTAGATCAGACTGAGCTGC 1560
DB |||||

QY 1561 AGCGAAGAGACCCCTAGGGAGGATCACACTGGATGAAGAGAGATGTGGAGAGTCCAGG 1620
DB |||||
QY 1561 AGCGAAGAGACCCCTAGGGAGGATCACACTGGATGAAGAGAGATGTGGAGAGTCCAGG 1620
DB |||||
QY 1621 GCAACCTAAGAGCCAGAGCCTTAAAGAGCAAGAGATAAAGGTGCTTCAAGGTGGCCAGG 1680
DB |||||
QY 1621 GCAACCTAAGAGCCAGAGCCTTAAAGAGCAAGAGATAAAGGTGCTTCAAGGTGGCCAGG 1680
DB |||||
QY 1681 CTGTGCAACAGAGGGTTCAGAGACTGTGGGTAGAGCTCAAGATGAAGATGATGCTCAGA 1740
DB |||||
QY 1681 CTGTGCAACAGAGGGTTCAGAGACTGTGGGTAGAGCTCAAGATGAAGATGATGCTCAGA 1740
DB |||||
QY 1741 ATGGCGGGGGGGGGATTCTGGGGGGGGGAGAGAGAGGTGAGAGAGCTTGAAGCAG 1800
DB |||||
QY 1741 ATGGCGGGGGGGGGATTCTGGGGGGGGGAGAGAGAGGTGAGAGAGCTTGAAGCAG 1800
DB |||||
QY 1801 AGAATCTGGAAAGCCTGGAAAACGATACATTAAGGGGAAAGAACCCAGGCTACTTTAGATG 1860
DB |||||
QY 1801 AGAATCTGGAAAGCCTGGAAAACGATACATTAAGGGGAAAGAACCCAGGCTACTTTAGATG 1860
DB |||||
QY 1861 TAAATCATGAAGACAGGGAGAAAGGAGCTGGAGAGAGTAGAAGGACCCCGGGCAGA 1920
DB |||||
QY 1861 TAAATCATGAAGACAGGGAGAAAGGAGCTGGAGAGAGTAGAAGGACCCCGGGCAGA 1920
DB |||||
QY 1921 CATGGAAGCAAGGACAAAGCCAGGTTGAGCGCTCCGTGAAATCAGCGCTCTGAAGCAGAG 1980
DB |||||
QY 1921 CATGGAAGCAAGGACAAAGCCAGGTTGAGCGCTCCGTGAAATCAGCGCTCTGAAGCAGAG 1980
DB |||||
QY 1981 CCCTGTATGAGCACCAAGAACAGCAGAGAGGCTAGGGTTAATGTGAGAGACAGGAAACAGAG 2040
DB |||||
QY 1981 CCCTGTATGAGCACCAAGAACAGCAGAGAGGCTAGGGTTAATGTGAGAGACAGGAAACAGAG 2040
DB |||||
QY 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAAACAAAGAAATGGTCTCTCAC 2100
DB |||||
QY 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAAACAAAGAAATGGTCTCTCAC 2100
DB |||||
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCATCAGACTGCAGGC 2160
DB |||||
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCATCAGACTGCAGGC 2160
DB |||||
QY 2161 TGGCTTGGGAGCGAGCTGGAAAGATGTGAGAGCCAGGGGAGACAAAGGGGGCTTAGGA 2220
DB |||||
QY 2161 TGGCTTGGGAGCGAGCTGGAAAGATGTGAGAGCCAGGGGAGACAAAGGGGGCTTAGGA 2220
DB |||||
QY 2221 AAGGAAGAGAGGGGCAACCAAGGCCACACAAAGAGGGGAGGCCAGAACTGAGTTAACTC 2280
DB |||||
QY 2221 AAGGAAGAGAGGGGCAACCAAGGCCACACAAAGAGGGGAGGCCAGAACTGAGTTAACTC 2280
DB |||||
QY 2281 CTTCTTGTGTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
DB |||||
QY 2281 CTTCTTGTGTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
DB |||||
QY 2341 GCCCCCCACTACCCATACCAAGTTTGGGCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG 2400
DB |||||
QY 2341 GCCCCCCACTACCCATACCAAGTTTGGGCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG 2400
DB |||||
QY 2401 CCTGGCTTTGTCTTCTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTGAAT 2460
DB |||||
QY 2401 CCTGGCTTTGTCTTCTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTTATCAT 2460
DB |||||
QY 2461 GCCCTCAAGAGCTTGAGAACCCAGGAGTGCATATTAGGCCATGGGCTAACCTGGAGCT 2520
DB |||||
QY 2461 GCCCTCAAGAGCTTGAGAACCCAGGAGTGCATATTAGGCCATGGGCTAACCTGGAGCT 2520
DB |||||
QY 2521 TGCACACAGGAGCTCAAGTCACTCCAGGAGACACAGCTGCAGACAGGTGGCTTTATCC 2580
DB |||||
QY 2521 TGCACACAGGAGCTCAAGTCACTCCAGGAGACACAGCTGCAGACAGGTGGCTTTATCC 2580
DB |||||
QY 2581 CCAAAGAGCAACCAATTTGGCATAGGTGGCTGCAAAATGGGAATGCAAGGTTGAATCAGGTC 2640
DB |||||
QY 2581 CCAAAGAGCAACCAATTTGGCATAGGTGGCTGCAAAATGGGAATGCAAGGTTGAATCAGGTC 2640
DB |||||
QY 2641 CCTTCAAGATACCTGATGCAAGACCTTAAGACCCCTGGAGAGAGGGGTATGCTCTCGCCC 2700
DB |||||

[illegible]

```
QY 4861 TCCCTTACGCTGCTCTTCCATCTTCTCCCTGGGATGCTCCTCCCGCTCTGGTTTATC 4920
Db 4862 TCCCTTACGCTGCTCTTCCATCTTCTCCCTGGGATGCTCCTCCCGCTCTGGTTTATC 4628
QY 4921 TTGGCTCTTCGCTTTCAGCAAGATTGGCTGTGCTGCCATCCATCTTCTCTACTGT 4980
Db 4629 TTGGCTCTTCGCTTTCAGCAAGATTGGCTGTGCTGCCATCCATCTTCTCTACTGT 4688
QY 4981 CTCGGTGCCTTGCCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 5040
Db 4689 CTCGGTGCCTTGCCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 4748
QY 5041 TTTTCTCCCTTCTCATTGTATTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5100
Db 4749 TTTTCTCCCTTCTCATTGTATTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4808
QY 5101 CTTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5160
Db 4809 CTTCCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4868
QY 5161 CTGTGTACAGTGTGAGATCACACCTGGGGTTCACACCTTATGTAAACAATCTTCCA 5220
Db 4869 CTGTGTACAGTGTGAGATCACACCTGGGGTTCACACCTTATGTAAACAATCTTCCA 4928
QY 5221 GTGAGCCACAGCTCAGTGTGCTGGGTGCTCTTACCTTCTTCTTCTTCTTCTTCT 5280
Db 4929 GTGAGCCACAGCTCAGTGTGCTGGGTGCTCTTACCTTCTTCTTCTTCTTCTTCT 4988
QY 5281 CTGTGTACAGTGTGAGATCACACCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5340
Db 4989 CTGTGTACAGTGTGAGATCACACCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5048
QY 5341 GCCTGTCTCTCTGTGTACAGTGTGCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5400
Db 5049 GCCTGTCTCTCTGTGTACAGTGTGCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5108
QY 5401 ATTCTTACAGTGTGAGATCACACCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5460
Db 5109 ATTCTTACAGTGTGAGATCACACCTGGGTGCTCTTACCTTCTTCTTCTTCTTCT 5168
QY 5461 TCACGATTTCTCCGGAAGTTCAGGCTTCCAGCCCTCTTCTTCTTCTTCTTCTTCT 5520
Db 5169 TCACGATTTCTCCGGAAGTTCAGGCTTCCAGCCCTCTTCTTCTTCTTCTTCTTCT 5228
QY 5521 CTCTTAGCAACCTCAGGCACTTACCCCATAGACTCTCTGACAGAGAGAGGCACT 5580
Db 5229 CTCTTAGCAACCTCAGGCACTTACCCCATAGACTCTCTGACAGAGAGAGGCACT 5288
QY 5581 TTACATGAGTCTTGGTGGAGAGCCATAGGCTACGGTGTAAAGAGAGGAGAGTGGT 5640
Db 5289 TTACATGAGTCTTGGTGGAGAGCCATAGGCTACGGTGTAAAGAGAGGAGAGTGGT 5348
QY 5641 GGTGTAGGAAGTTCAGGATTCACATAGAGCTTAGCCACACACAGAGAGATGACAGACA 5700
Db 5349 GGTGTAGGAAGTTCAGGATTCACATAGAGCTTAGCCACACACAGAGAGATGACAGACA 5408
QY 5701 TCCCTCTCTATCTCCCTTACAGATTTTGGATGAC 5735
Db 5409 TCCCTCTCTATCTCCCTTACAGATTTTGGATGAC 5443
```

RESULT 3

```
US-10-332-966-2
; Sequence 2, Application US/10332966
; Publication No. US20030188324A1
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, Koji
; APPLICANT: KAWASE, Yosuke
; APPLICANT: SUZUKI, Hiroshi
; TITLE OF INVENTION: p300 TRANSGENIC ANIMAL
; FILE REFERENCE: 382.1040
; CURRENT APPLICATION NUMBER: US/10/332,966
```

```
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/JP01/06086
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: JP2000-215143
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(5443)
US-10-332-966-2
```

Query Match 88.0%; Score 5045; DB 16; Length 5443;

Best Local Similarity 93.9%; Pred. No. 0;

Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

```
QY 1 GGATCCTCAGAGTACACAAAGGCTCTCCACCACAGGTGCCCTAGTCTCAATTTCACT 60
Db 1 GGATCCTCAGAGTACACAAAGGCTCTCCACCACAGGTGCCCTAGTCTCAATTTCACT 60
QY 61 TTCCATGCTTTGTTCTCAATGCTGGGCTCCCCAGAGCTAATTTGGACTTTGTTTAT 120
Db 61 TTCCATGCTTTGTTCTCAATGCTGGGCTCCCCAGAGCTAATTTGGACTTTGTTTAT 120
QY 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTTGTGTACCCAGCTTAAGGGTCCCGTGA 180
Db 121 TTCAAAAGGGCTGAATGAGGAGTAGATCTTTGTGTACCCAGCTTAAGGGTCCCGTGA 180
QY 181 AGCCCTCAGACTGGAGCTTTGCAAGAGCCCTTTAGTGGAGAGCAGATAAAGCAATTT 240
Db 181 AGCCCTCAGACTGGAGCTTTGCAAGAGCCCTTTAGTGGAGAGCAGATAAAGCAATTT 240
QY 241 TCCTTAAAGCCAAATCTCGCTCTAGACTCTTTCTTCTGACCTCGGTCCTGGGCTCT 300
Db 241 TCCTTAAAGCCAAATCTCGCTCTAGACTCTTTCTTCTGACCTCGGTCCTGGGCTCT 300
QY 301 AGGGTGGGAGGTGGGGCTTGGAAAGAGAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
Db 301 AGGGTGGGAGGTGGGGCTTGGAAAGAGAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
QY 361 GGCCCTGTGAAGTTCGGAGCCCTTCCCTGTAAGACTGGCTCATAGATCTCTCTCAGCC 420
Db 361 GGCCCTGTGAAGTTCGGAGCCCTTCCCTGTAAGACTGGCTCATAGATCTCTCTCAGCC 420
QY 421 AAACATAGCAAGAGTGTATCTCTTGTGACTTCCCGAGCCCGAGTACCTGTGAGTT 480
Db 421 AAACATAGCAAGAGTGTATCTCTTGTGACTTCCCGAGCCCGAGTACCTGTGAGTT 480
QY 481 GAAACAGAGATTTAGAGAGGCTCTGAACTCACTGAACTCTGAACTCATCCACCAAGCA 540
Db 481 GAAACAGAGATTTAGAGAGGCTCTGAACTCACTGAACTCATCCACCAAGCA 540
QY 541 AGCAGCTAGGTGCCACTGCTAGTATCTTACGCTGATATATATGAGAGCTGGGCCAC 600
Db 541 AGCAGCTAGGTGCCACTGCTAGTATCTTACGCTGATATATATGAGAGCTGGGCCAC 600
QY 601 AGNAGTCTGGGTGTAGGAACTGACAGTGTATTTTTCAGTGGCAAGGTATGACCCCC 660
Db 601 AGNAGTCTGGGTGTAGGAACTGACAGTGTATTTTTCAGTGGCAAGGTATGACCCCC 660
QY 661 TCAGCAGATGTAGTAATGTCCCTCTAGATCCCATCCAGGAGGTCTCTTAAGAGGACATG 720
Db 661 TCAGCAGATGTAGTAATGTCCCTCTAGATCCCATCCAGGAGGTCTCTTAAGAGGACATG 720
QY 721 GGATGAGAGATGTAGTCTATGTGGCAATCCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780
Db 721 GGATGAGAGATGTAGTCTATGTGGCAATCCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780
QY 781 TTCCACTAGCCAGGAGGACAGTAACTTAGCTATCTTTCTTCTCTCCCATCTCTCCAG 840
```

| | | | |
|----|------|---|------|
| Db | 781 | TTCCACTTAGCCAGGAGACAGTAACTTAGCCTATCTTTCTCTCCCTCCCTCCAG | 840 |
| Qy | 841 | GACACACCCCTGGTCTCAGTATTCATTTCTTCTTCAAGTCCCTCTGTGACTTCAT | 900 |
| Db | 841 | GACACACCCCTGGTCTCAGTATTCATTTCTTCTTCAAGTCCCTCTGTGACTTCAT | 900 |
| Qy | 901 | TTGCAAGGCTTTTGACCTCTGCAGCTCTGGAAGATAGAGTTTGGCCCTTAGGTGGCAA | 960 |
| Db | 901 | TTGCAAGGCTTTTGACCTCTGCAGCTCTGGAAGATAGAGTTTGGCCCTTAGGTGGCAA | 960 |
| Qy | 961 | GCCATCTCAAGAAAGCAGACAAACAGGGGACAGATTTTGGAGGATCAGGAACTAAA | 1020 |
| Db | 961 | GCCATCTCAAGAAAGCAGACAAACAGGGGACAGATTTTGGAGGATCAGGAACTAAA | 1020 |
| Qy | 1021 | TCATCGGGGGCTGGGGGTAGAAAAAGAGTGAAGTCCGCTCAGCTAAGCCAAGC | 1080 |
| Db | 1021 | TCATCGGGGGCTGGGGGTAGAAAAAGAGTGAAGTCCGCTCAGCTAAGCCAAGC | 1080 |
| Qy | 1081 | TAGTCCCCGAGATCTCTGCCACAGCTGGGCTCTCGGGGTAGCTTTAGGAATGTGGTCT | 1140 |
| Db | 1081 | TAGTCCCCGAGATCTCTGCCACAGCTGGGCTCTCGGGGTAGCTTTAGGAATGTGGTCT | 1140 |
| Qy | 1141 | TGAAAGCAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCCACTACAGAC | 1200 |
| Db | 1141 | TGAAAGCAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCCACTACAGAC | 1200 |
| Qy | 1201 | ACACTCGTGTGGCCAGACTCTGTTCACAGCCCTCTGTCTGACCACTGAGCTAG | 1260 |
| Db | 1201 | ACACTCGTGTGGCCAGACTCTGTTCACAGCCCTCTGTCTGACCACTGAGCTAG | 1260 |
| Qy | 1261 | GCAACAGAGATGGGCTGTGTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG | 1320 |
| Db | 1261 | GCAACAGAGATGGGCTGTGTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG | 1320 |
| Qy | 1321 | CAGGGAGGAGAAACAGAGAACGAAATTAAGGAAGGAAGAAAGGCCAGTCAATCAGA | 1380 |
| Db | 1321 | CAGGGAGGAGAAACAGAGAACGAAATTAAGGAAGGAAGAAAGGCCAGTCAATCAGA | 1380 |
| Qy | 1381 | TGCAGTCAAGAGATGGGAAGCCAAACACAGCTTGAAGAGGAAAGGAAAGGGAG | 1440 |
| Db | 1381 | TGCAGTCAAGAGATGGGAAGCCAAACACAGCTTGAAGAGGAAAGGAAAGGGAG | 1440 |
| Qy | 1441 | AGATTCTGGCATTAAGAGGCCACAGAAAGAGAGCCAGCCGCCCAAGTCTCTCTTT | 1500 |
| Db | 1441 | AGATTCTGGCATTAAGAGGCCACAGAAAGAGAGCCAGCCGCCCAAGTCTCTCTTT | 1500 |
| Qy | 1501 | ATACTCTCATCCGCTCTCCCAATTAAGCCCACTCTTTCTTCTAGATCAGACTGAGCTGC | 1560 |
| Db | 1501 | ATACTCTCATCCGCTCTCCCAATTAAGCCCACTCTTTCTTCTAGATCAGACTGAGCTGC | 1560 |
| Qy | 1561 | AGCGAAGAGACCCGTAGGAGGATCACAAGTGAAGAGATGTGTGGAGATGTCAGG | 1620 |
| Db | 1561 | AGCGAAGAGACCCGTAGGAGGATCACAAGTGAAGAGATGTGTGGAGATGTCAGG | 1620 |
| Qy | 1621 | GCAACTAAGAGCCAGAGCTTAAAGAGCAAGAGATAAGTGTTCAAAGTGGCCAGG | 1680 |
| Db | 1621 | GCAACTAAGAGCCAGAGCTTAAAGAGCAAGAGATAAGTGTTCAAAGTGGCCAGG | 1680 |
| Qy | 1681 | CTGTGCACACAGAGGCTGAGGACTGTGGTGTAGAGCTTCAAGATAAGGATGATCTCAGA | 1740 |
| Db | 1681 | CTGTGCACACAGAGGCTGAGGACTGTGGTGTAGAGCTTCAAGATAAGGATGATCTCAGA | 1740 |
| Qy | 1741 | ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAAAGGTGAGAGAGCTGGAAACAG | 1800 |
| Db | 1741 | ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAAAGGTGAGAGAGCTGGAAACAG | 1800 |
| Qy | 1801 | AGAATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGAACCCAGGCTACTTTAGATG | 1860 |
| Db | 1801 | AGAATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGAACCCAGGCTACTTTAGATG | 1860 |
| Qy | 1861 | TAAATCATGAAGACAGGAGAGGAGCTGGAGAGATGAAGAGACCCCGGGGCAAGA | 1920 |
| Db | 1861 | TAAATCATGAAGACAGGAGAGGAGCTGGAGAGATGAAGAGACCCCGGGGCAAGA | 1920 |

| | | | |
|----|------|--|------|
| Db | 1861 | TAAATCATGAAGACAGGAGAGGAGAGCTGGAGAGATGAAGAGACCCCGGGGCAAGA | 1920 |
| Qy | 1921 | CATGGAAGCAGGAGCAAGCCAGGTTGAGCGCTCGTGAATCAGCCTCTGAAGCAGAG | 1980 |
| Db | 1921 | CATGGAAGCAGGAGCAAGCCAGGTTGAGCGCTCGTGAATCAGCCTCTGAAGCAGAG | 1980 |
| Qy | 1981 | CCCTGGTATGAGCACAGAAACAGCAGAGGCTAGGGTTAATGTTCGAGACAGGAAACAGAG | 2040 |
| Db | 1981 | CCCTGGTATGAGCACAGAAACAGCAGAGGCTAGGGTTAATGTTCGAGACAGGAAACAGAG | 2040 |
| Qy | 2041 | GTAGACACAGAAACAGACAGAGACGCGGAGCCAGSTAAACAAAGAAATGGTCTCTCAC | 2100 |
| Db | 2041 | GTAGACACAGAAACAGACAGAGACGCGGAGCCAGSTAAACAAAGAAATGGTCTCTCAC | 2100 |
| Qy | 2101 | CTGTGGCCAGAGCGCTCATCTGTCTCCACATCTCTAGAAATGTTTCATCAGACTGAGGCG | 2160 |
| Db | 2101 | CTGTGGCCAGAGCGCTCATCTGTCTCCACATCTCTAGAAATGTTTCATCAGACTGAGGCG | 2160 |
| Qy | 2161 | TGGCTTGGGAGGCGAGCTGGAAAGATATGTGAGAGCCAGGGGAGACAAAGGGGCTTAGGA | 2220 |
| Db | 2161 | TGGCTTGGGAGGCGAGCTGGAAAGATATGTGAGAGCCAGGGGAGACAAAGGGGCTTAGGA | 2220 |
| Qy | 2221 | AAGGAAGAGAGGCGAAACCAAGGCCACACAAAGAGGCGAGAGCCAGAACTGAGTTAACTC | 2280 |
| Db | 2221 | AAGGAAGAGAGGCGAAACCAAGGCCACACAAAGAGGCGAGAGCCAGAACTGAGTTAACTC | 2280 |
| Qy | 2281 | CTTCTTGTGTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCCACCATCCCTCATGA | 2340 |
| Db | 2281 | CTTCTTGTGTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCCACCATCCCTCATGA | 2340 |
| Qy | 2341 | GCCCCCACTACCCATACCAAGTTTGGCTGAGTGGCAATTTAGGTTCCCTGAGGACAGAG | 2400 |
| Db | 2341 | GCCCCCACTACCCATACCAAGTTTGGCTGAGTGGCAATTTAGGTTCCCTGAGGACAGAG | 2400 |
| Qy | 2401 | CCTGGCTTTGTCTCTTTGGACCTGACCCAAAGCTGACCAATGTTCTCAGTACCTTTGAAT | 2460 |
| Db | 2401 | CCTGGCTTTGTCTCTTTGGACCTGACCCAAAGCTGACCAATGTTCTCAGTACCTTTGAAT | 2460 |
| Qy | 2461 | GCCCTCAAGAGCTTTGAGAACCCAGGACGTGACATATTAGGCCATGGCTTAAACCTGGAGCT | 2520 |
| Db | 2461 | GCCCTCAAGAGCTTTGAGAACCCAGGACGTGACATATTAGGCCATGGCTTAAACCTGGAGCT | 2520 |
| Qy | 2521 | TGCAACAGGAGCTCAAGTGACCTCAGGGACACACAGCTGACAGAGGCTGCTTTATCC | 2580 |
| Db | 2521 | TGCAACAGGAGCTCAAGTGACCTCAGGGACACACAGCTGACAGAGGCTGCTTTATCC | 2580 |
| Qy | 2581 | CCAAAGAGCAACCAATTTGGCATAGGTGGCTGCAAAATGGGAATGCAAGGTTGAATCAGGTC | 2640 |
| Db | 2581 | CCAAAGAGCAACCAATTTGGCATAGGTGGCTGCAAAATGGGAATGCAAGGTTGAATCAGGTC | 2640 |
| Qy | 2641 | CCTTCAAGATACCTGCAATGCAAGACCTTAAGACCCCTGGAGAGAGGGTATGCTCTGCCC | 2700 |
| Db | 2641 | CCTTCAAGATACCTGCAATGCAAGACCTTAAGACCCCTGGAGAGAGGGTATGCTCTGCCC | 2700 |
| Qy | 2701 | CCACCCACCAATAGGGGAGTGAACTATCTAGGGGCTGGCGACCTTGGGAGACACAC | 2760 |
| Db | 2701 | CCACCCACCAATAGGGGAGTGAACTATCTAGGGGCTGGCGACCTTGGGAGACACAC | 2760 |
| Qy | 2761 | ATTAAGAGTGTGAGCCAGAAAACTGACCCGCTGTGTCTGCTGCCACCTCCAC | 2820 |
| Db | 2761 | ATTAAGAGTGTGAGCCAGAAAACTGACCCGCTGTGTCTGCTGCCACCTCCAC | 2820 |
| Qy | 2821 | TCTAGAGCTATTTAGAGGTTGACGTAGATAGGTGGAGCTGGTAGCAGGAGAGTGT | 2880 |
| Db | 2821 | TCTAGAGCTATTTAGAGGTTGACGTAGATAGGTGGAGCTGGTAGCAGGAGAGTGT | 2880 |
| Qy | 2881 | TCCTGGGTGTGAGGCTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTCTGAA | 2940 |
| Db | 2881 | TCCTGGGTGTGAGGCTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTCTGAA | 2940 |
| Qy | 2941 | CACAAATGTCTTATGATTAACAGGATGACTGCTTAAAGACCCCAACATCTACGACCTC | 3000 |
| Db | 2941 | CACAAATGTCTTATGATTAACAGGATGACTGCTTAAAGACCCCAACATCTACGACCTC | 3000 |


```
Db      ||||||| 4869 CTGTGTCAAGTGTGAGATCACACCTGGGGTTCCCAACCTTATGTAAACAATCTTCCA 4928
Qy      ||||||| 5221 GTGAGCCACAGCTCAGTGTGGTGTCTCTTACCTTCTCACCCTCGCTGTGC 5280
Db      ||||||| 4929 GTGAGCCACAGCTCAGTGTGGTGTCTCTTACCTTCTCACCCTCGCTGTGC 4988
Qy      ||||||| 5281 CTGTTCCATCTGTGTGAGGATCTCTAGATTTGGTCTTCCAGAGCTCTGTCTTCTTCTT 5340
Db      ||||||| 4989 CTGTTCCATCTGTGTGAGGATCTCTAGATTTGGTCTTCCAGAGCTCTGTCTTCTTCTT 5048
Qy      ||||||| 5341 GCCTGTTCTCTCTCTGTGTCAGTGTGGCCACTGTGTGTGCTCTGTGTGTCAGTGTGTGTCAC 5400
Db      ||||||| 5049 GCCTGTTCTCTCTCTGTGTCAGTGTGGCCACTGTGTGTGCTCTGTGTGTCAGTGTGTGTCAC 5108
Qy      ||||||| 5401 ATTCTTCAGGATTTCTCTGAAAGTTAAACAGGTGAGATGTTCCCTGTGACAGCAGA 5460
Db      ||||||| 5109 ATTCTTCAGGATTTCTCTGAAAGTTAAACAGGTGAGATGTTCCCTGTGACAGCAGA 5168
Qy      ||||||| 5461 TCAGATTTCTCCGGAAGTCAGGCTTCCAGGCCCTCTCTTTCTCTGCCCAGCTGCCGGCA 5520
Db      ||||||| 5169 TCAGATTTCTCCGGAAGTCAGGCTTCCAGGCCCTCTCTTTCTCTGCCCAGCTGCCGGCA 5228
Qy      ||||||| 5521 CTCTTAGCAACCTCAGGCACTCTTAACCCACATAGACCTCTGACAGAAAGCAGGCACT 5580
Db      ||||||| 5229 CTCTTAGCAACCTCAGGCACTCTTAACCCACATAGACCTCTGACAGAAAGCAGGCACT 5288
Qy      ||||||| 5581 TTACATGAGTCTCTGTGGGAGCCATAGCTACGGTGTAAAGAGCAGGGAAGTGGT 5640
Db      ||||||| 5289 TTACATGAGTCTCTGTGGGAGCCATAGCTACGGTGTAAAGAGCAGGGAAGTGGT 5348
Qy      ||||||| 5641 GGTGTAGGAAGTCAGGATTTCAATAGAGCCTAGCCACACAGAGAAATGACAGACAGA 5700
Db      ||||||| 5349 GGTGTAGGAAGTCAGGATTTCAATAGAGCCTAGCCACACAGAGAAATGACAGACAGA 5408
Qy      ||||||| 5701 TCCCTCTCTATCTCCCCCATAGAGTTTGAGTCGAC 5735
Db      ||||||| 5409 TCCCTCTCTATCTCCCCCATAGAGTTTGAGTCGAC 5443
```

RESULT 4

```
US-10-944-375-15
; Sequence 15, Application US/10944375
; Publication No. US20050066381A1
; GENERAL INFORMATION:
; APPLICANT: Cincinnati Children's Hospital Medical Center
; APPLICANT: University of Cincinnati
; APPLICANT: Molkenin, Jeffery D
; APPLICANT: Kranias, Evangelia G
; TITLE OF INVENTION: REGULATION OF CARDIAC CONTRACTILITY AND HEART FAILURE PROPENSITY
; FILE REFERENCE: 9761M#L
; CURRENT APPLICATION NUMBER: US/10/944,375
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: US 60/503,853
; PRIOR FILING DATE: 2003-09-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-944-375-15
```

```
Query Match      88.0%; Score 5045; DB 19; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

Qy      1  GGATCTCGAAGGTCAACAAGGGTCTCCACCCACAGGTGCCCTAGTCTCAATTTTCAGT 60
Db      1  GGATCTCGAAGGTCAACAAGGGTCTCCACCCACAGGTGCCCTAGTCTCAATTTTCAGT 60
Qy      61  TTCCATGCTTGTCTTCAAAATGCTGGCTCCCAAGAGCTAATTTGGACTTTGTTTAT 120
```

```
Db      61  TTCCATGCTTGTCTTCAAAATGCTGGCTCCCAAGAGCTAATTTGGACTTTGTTTAT 120
Qy      121  TTCAAAAGGGCTCGAATGAGGAGTAGATCTTGTCTACCCAGCTCTAAGGGTGGCCGTGA 180
Db      121  TTCAAAAGGGCTCGAATGAGGAGTAGATCTTGTCTACCCAGCTCTAAGGGTGGCCGTGA 180
Qy      181  AGCCCTCAGACCTTGAGCCCTTTGCAACAGCCCTTTAGTGGAGCAGCAATAAAGCAATTT 240
Db      181  AGCCCTCAGACCTTGAGCCCTTTGCAACAGCCCTTTAGTGGAGCAGCAATAAAGCAATTT 240
Qy      241  TCCTTAAAGCCAAATCTCTGCTCTAGACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db      241  TCCTTAAAGCCAAATCTCTGCTCTAGACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Qy      301  AGGTGCGGAGGTGGGGCTTTGGAAGAAAGGTGGGAAAGTGGCAAAAGCCGATCCCTTAG 360
Db      301  AGGTGCGGAGGTGGGGCTTTGGAAGAAAGGTGGGAAAGTGGCAAAAGCCGATCCCTTAG 360
Qy      361  GGCCCTGTGAAGTTTCGAGGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCTCCAGCC 420
Db      361  GGCCCTGTGAAGTTTCGAGGCTTCCCTGTACAGCACTGGCTCATAGATCTCTCTCCAGCC 420
Qy      421  AAACATAGCAAGAGTGTATACCTCTTTGTGACTTCCCGAGCCGAGTACCTGTGAGGTT 480
Db      421  AAACATAGCAAGAGTGTATACCTCTTTGTGACTTCCCGAGCCGAGTACCTGTGAGGTT 480
Qy      481  GAAACAGGATTTAGAGAGGCTCTGAACTCACCTGAACTCTGAAAGTCTATGAGAGTCTGGGCCAC 540
Db      481  GAAACAGGATTTAGAGAGGCTCTGAACTCACCTGAACTCTGAAAGTCTATGAGAGTCTGGGCCAC 540
Qy      541  AGCACTTGTAGTGTGCTGTAGTGTATCTCTAGTGTATCTCTGCTGATATATGAGAGTCTGGGCCAC 600
Db      541  AGCACTTGTAGTGTGCTGTAGTGTATCTCTAGTGTATCTCTGCTGATATATGAGAGTCTGGGCCAC 600
Qy      601  AGAAGTCTGGGGTGTAGGAACTGACCAAGTACTTTTCAGTGGGAAAGGTATGACCCCC 660
Db      601  AGAAGTCTGGGGTGTAGGAACTGACCAAGTACTTTTCAGTGGGAAAGGTATGACCCCC 660
Qy      661  TCAGCAGATGTAGTAATGTCCCTTTAGTATCCATCCAGGAGGCTCTCTAAGAGGACATG 720
Db      661  TCAGCAGATGTAGTAATGTCCCTTTAGTATCCATCCAGGAGGCTCTCTAAGAGGACATG 720
Qy      721  GGATGAGAGATGTAGTCAATGTGGCAATCCAAACACAGCTATCCACAGTGTCCCTTGGCCC 780
Db      721  GGATGAGAGATGTAGTCAATGTGGCAATCCAAACACAGCTATCCACAGTGTCCCTTGGCCC 780
Qy      781  TTCCATTTAGCAGGAGGACAGATTAACCTTAGCCCTATCTTTCTCTCCCAATCTCTCCAG 840
Db      781  TTCCATTTAGCAGGAGGACAGATTAACCTTAGCCCTATCTTTCTCTCCCAATCTCTCCAG 840
Qy      841  GACACACCTCTGTCTGAGTATTTCAATTTCTTCTTCACTGCTCCCTCTGTGACTTCCAT 900
Db      841  GACACACCTCTGTCTGAGTATTTCAATTTCTTCTTCACTGCTCCCTCTGTGACTTCCAT 900
Qy      901  TTGCAAGGCTTTTGACCTCTGACGCTGTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960
Db      901  TTGCAAGGCTTTTGACCTCTGACGCTGTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960
Qy      961  GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGATTTTGGAGAGGATCAGGAATCTAAA 1020
Db      961  GCCATCTCAAGAGAAAGCAGACAAACAGGGGACAGATTTTGGAGAGGATCAGGAATCTAAA 1020
Qy      1021  TCACCTGGGGGCTGGGGGTAGAAAAGAGTGTAGTCCGCTCCAGCTTAAGCCAAAGC 1080
Db      1021  TCACCTGGGGGCTGGGGGTAGAAAAGAGTGTAGTCCGCTCCAGCTTAAGCCAAAGC 1080
Qy      1081  TAGTCTCCCGAGATCTCTGCGCACAGCTGGGTGTCTCGGGGTAGCTTTTAGGAATGTGGGTC 1140
Db      1081  TAGTCTCCCGAGATCTCTGCGCACAGCTGGGTGTCTCGGGGTAGCTTTTAGGAATGTGGGTC 1140
Qy      1141  TGAAGAACAATGGAGATGGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
Db      1141  TGAAGAACAATGGAGATGGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
```

QY 1201 ACATCGTGTGTGGCCAGACTCTGTCTCAACAGCCCTCTGTGTCTGTGACCACTGAGCTAG 1260
DB |||||
QY 1201 ACATCGTGTGTGGCCAGACTCTGTCTCAACAGCCCTCTGTGTCTGTGACCACTGAGCTAG 1260
DB |||||
QY 1261 GCAACAGAGATGGGCCCTGTGTGAGGATGAAGAGTGTGTGTTACCAATAGCAAAACAG 1320
DB |||||
QY 1261 GCAACAGAGATGGGCCCTGTGTGAGGATGAAGAGTGTGTGTTACCAATAGCAAAACAG 1320
DB |||||
QY 1321 CAGGGAGGAGAACAGAGAACGAATTAAGGAAGGAAGGAAGGCGCAGTCAATCAGA 1380
DB |||||
QY 1321 CAGGGAGGAGAACAGAGAACGAATTAAGGAAGGAAGGAAGGCGCAGTCAATCAGA 1380
DB |||||
QY 1381 TGCAGTCAGAGAGATGGGAAGCCAAACACAGCTTTGAGCAGAGAAACAGAAAGGGAG 1440
DB |||||
QY 1381 TGCAGTCAGAGAGATGGGAAGCCAAACACAGCTTTGAGCAGAGAAACAGAAAGGGAG 1440
DB |||||
QY 1441 AGATTCTGGGCATAGAGGAGCCACAGAAAGAGAGCCAGGCCCCCAAGTCTCTCTTT 1500
DB |||||
QY 1441 AGATTCTGGGCATAGAGGAGCCACAGAAAGAGAGCCAGGCCCCCAAGTCTCTCTTT 1500
DB |||||
QY 1501 ATACCCCTCATCCCGTCTCCCAATTAAGGCCACTCTTCTTCTAGATCAGACCTGAGCTGC 1560
DB |||||
QY 1501 ATACCCCTCATCCCGTCTCCCAATTAAGGCCACTCTTCTTCTAGATCAGACCTGAGCTGC 1560
DB |||||
QY 1561 AGCGAAGAGACCCGTAGGAGGATCACACTGGATGAAGAGATGTGTGGGAAGTCCAGG 1620
DB |||||
QY 1561 AGCGAAGAGACCCGTAGGAGGATCACACTGGATGAAGAGATGTGTGGGAAGTCCAGG 1620
DB |||||
QY 1621 GCAACCTTAAGAGCCAGAGCCTAAAGAGCAAGAGATAAAGGTGCTTCAAAAGTGGCCAGG 1680
DB |||||
QY 1621 GCAACCTTAAGAGCCAGAGCCTAAAGAGCAAGAGATAAAGGTGCTTCAAAAGTGGCCAGG 1680
DB |||||
QY 1681 CTGTGCAACAGAGGGTCAAGGATGTGTGTGTGTAGAGCCTCAAGATTAAGATGATGCTCAGA 1740
DB |||||
QY 1681 CTGTGCAACAGAGGGTCAAGGATGTGTGTGTGTGTAGAGCCTCAAGATTAAGATGATGCTCAGA 1740
DB |||||
QY 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGGTGAGAGAGGCTTGAAACAG 1800
DB |||||
QY 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGGTGAGAGAGGCTTGAAACAG 1800
DB |||||
QY 1801 AGAATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGAACCCAGGCTACCTTTAGATG 1860
DB |||||
QY 1801 AGAATCTGGAAGCGCTGGAACGATACCATTAAGGGAAGAACCCAGGCTACCTTTAGATG 1860
DB |||||
QY 1861 TAAATCATGAAGACAGGGAGAGGAAGAGTGGAGAGAGTGAAGGACCCCGGGGCAAGA 1920
DB |||||
QY 1861 TAAATCATGAAGACAGGGAGAGGAAGAGTGGAGAGAGTGAAGGACCCCGGGGCAAGA 1920
DB |||||
QY 1921 CATGGAAGCAAGGACAGCCAGGTTGAGGGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980
DB |||||
QY 1921 CATGGAAGCAAGGACAGCCAGGTTGAGGGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980
DB |||||
QY 1981 CCCTGGTATGACACAGAAACAGCAGAGGCTAGGGTTAATGTGAGAGACAGGGAACAGAG 2040
DB |||||
QY 1981 CCCTGGTATGACACAGAAACAGCAGAGGCTAGGGTTAATGTGAGAGACAGGGAACAGAG 2040
DB |||||
QY 2041 GTAGACACAGGAACAGACAGAGACGGGGGAGCCAGGTAAACAAAGGAATGGTCTCTCAC 2100
DB |||||
QY 2041 GTAGACACAGGAACAGACAGAGACGGGGGAGCCAGGTAAACAAAGGAATGGTCTCTCAC 2100
DB |||||
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCAATCAGACTGAGGGC 2160
DB |||||
QY 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATCTCTAGAAATGTTTCAATCAGACTGAGGGC 2160
DB |||||
QY 2161 TGGCTTGGGGGCGAGCTGGAAGAGTATGTGAGCCAGGGGAGACAGGGGGCTTAGGA 2220
DB |||||
QY 2161 TGGCTTGGGGGCGAGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAGGGGGCTTAGGA 2220
DB |||||
QY 2221 AAGGAAGAGAGGGCAACAGGCCACACAGAGGGCAGAGCCAGAACTGAGTTAACTC 2280
DB |||||
QY 2221 AAGGAAGAGAGGGCAACAGGCCACACAGAGGGCAGAGCCAGAACTGAGTTAACTC 2280
DB |||||

QY 2281 CTTCTTGTTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
DB |||||
QY 2281 CTTCTTGTTCATCTTCCATAGGAGGAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
DB |||||
QY 2341 GCGGCCACTACCCATACCAAGTTTGGCTGAGTGGCATTTAGGTTCCCTGAGGACAGAG 2400
DB |||||
QY 2341 GCGGCCACTACCCATACCAAGTTTGGCTGAGTGGCATTTAGGTTCCCTGAGGACAGAG 2400
DB |||||
QY 2401 CTTGGCTTGTCTCTTGGACCTGACCCCAAGCTGACCAATGTTCTCAGTACCTTATCAT 2460
DB |||||
QY 2401 CTTGGCTTGTCTCTTGGACCTGACCCCAAGCTGACCAATGTTCTCAGTACCTTATCAT 2460
DB |||||
QY 2461 GCGCTCAAGAGCTTGAGAACCCAGGCACTGACATATTAGGCCATGGGCTAAACCTCGAGCT 2520
DB |||||
QY 2461 GCGCTCAAGAGCTTGAGAACCCAGGCACTGACATATTAGGCCATGGGCTAAACCTCGAGCT 2520
DB |||||
QY 2521 TGCACACAGGAGCTCAAGTGACTTCCAGGACACAGCTGACAGACAGGTGGCTTTTATCC 2580
DB |||||
QY 2521 TGCACACAGGAGCTCAAGTGACTTCCAGGACACAGCTGACAGACAGGTGGCTTTTATCC 2580
DB |||||
QY 2581 CCAAGAGCAACCAATTTGGCATAGTGGCTGCAATGGGAATGCAAGGTTGAATCAGGTC 2640
DB |||||
QY 2581 CCAAGAGCAACCAATTTGGCATAGTGGCTGCAATGGGAATGCAAGGTTGAATCAGGTC 2640
DB |||||
QY 2641 CTTCAAGAACTACTGCAATGCAAGACCTTAAGACCCCTGAGAGAGGGGTATGCTCTGCC 2700
DB |||||
QY 2641 CTTCAAGAACTACTGCAATGCAAGACCTTAAGACCCCTGAGAGAGGGGTATGCTCTGCC 2700
DB |||||
QY 2701 CCACCCACATTAAGGGAGTGAATCTTCTAGGGGGCTGGCGACCTTGGGAGACACAC 2760
DB |||||
QY 2701 CCACCCACATTAAGGGAGTGAATCTTCTAGGGGGCTGGCGACCTTGGGAGACACAC 2760
DB |||||
QY 2761 ATTACTGAGAGTGTGAGCCCAAGAAAACCTGACCGCCCTGTGCTGCGCCACCTCCACAC 2820
DB |||||
QY 2761 ATTACTGAGAGTGTGAGCCCAAGAAAACCTGACCGCCCTGTGCTGCGCCACCTCCACAC 2820
DB |||||
QY 2821 TCTAGAGCTATTTAGAGAGGTGACAGTAGATAGGTGGGAGCTGTGTAGCAGGAGAGTGT 2880
DB |||||
QY 2821 TCTAGAGCTATTTAGAGAGGTGACAGTAGATAGGTGGGAGCTGTGTAGCAGGAGAGTGT 2880
DB |||||
QY 2881 TCCTGGGTGTCAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
DB |||||
QY 2881 TCCTGGGTGTCAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
DB |||||
QY 2941 CACAAATGCTACTTATTAAGAGGCTGCTGCTTAAAGACCCCAACATCTACGACCTC 3000
DB |||||
QY 2941 CACAAATGCTACTTATTAAGAGGCTGCTGCTTAAAGACCCCAACATCTACGACCTC 3000
DB |||||
QY 3001 TGAAAGACAGCCCTGAGGAGCAGGGGTGCTCTGAGCCTTGGGTGCTTGTATGTTG 3060
DB |||||
QY 3001 TGAAAGACAGCCCTGAGGAGCAGGGGTGCTCTGAGCCTTGGGTGCTTGTATGTTG 3060
DB |||||
QY 3061 CCACAAAGAGGAGGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCATTG 3120
DB |||||
QY 3061 CCACAAAGAGGAGGATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAGGGCATTG 3120
DB |||||
QY 3121 GGAAGGGTCTAGTCTGACAGCCCTATCCATGGAATCTGAGCCTGAGGCTGGGGCCAACTGTTG 3180
DB |||||
QY 3121 GGAAGGGTCTAGTCTGACAGCCCTATCCATGGAATCTGAGCCTGAGGCTGGGGCCAACTGTTG 3180
DB |||||
QY 3181 TAAATCTCTGGGCTGCGCAGGCAATTAAGCAGACCTGCACTCTCTGCGAGCCTGGGGA 3240
DB |||||
QY 3181 TAAATCTCTGGGCTGCGCAGGCAATTAAGCAGACCTGCACTCTCTGCGAGCCTGGGGA 3240
DB |||||
QY 3241 GCGGAGAGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAAACCT 3300
DB |||||
QY 3241 GCGGAGAGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCCAGGATTAAACCT 3300
DB |||||
QY 3301 CTGGCTTCCCTTCCACCTTCCATCAGGAGTGGAGGGTTGACAGAGGGGTAAGAAA 3360
DB |||||
QY 3301 CTGGCTTCCCTTCCACCTTCCATCAGGAGTGGAGGGTTGACAGAGGGGTAAGAAA 3360
DB |||||
QY 3361 CCTACATGTCCAAACATCATGTGTGACGATATATGGATCAGTATGTGTAGAGGCAAGAA 3420


```
OY 3453 AAAGTCTGTGTCATGTGTGTGTGTCGACTGAAAAACGGGATGGCTGTGAGCTGTTC 3512
Db 3473 -----TCTGATGTGTCTGTGTC 3492
OY 3513 GTTCTGTGCGTGAGGTTACAGACTGCGAGGTTTGTGTAAATTTCCCAAGCAAGTGG 3572
Db 3493 AGTCTGCACGGGAGGCTGTGCGCAAGGGTCTGTGCGCAGGTGTGCTGAGGACGCTG 3552
OY 3573 GTGAATCCCTTCCATGGTTTAAAGAGATTGGATGATGCGCCTGCATCTCAAGGACCATGA 3632
Db 3553 GAGAAATTCCTTCCATCGGTGATGGTTTGCATGAGGGCCACTGTATGGGAGACCCGG 3612
OY 3633 AAATAGAAATGACACTCTATATGTGTCTTAAGCTAAGTAGCAAGGTCCTTTGAGGAC 3692
Db 3613 GAAGTGTGCGGCTCCAGATCAATCCCAAGCAAGCAGCCTTGAGTCTCTGTGACACA 3672
OY 3693 CCTGTCTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGATGAGGAGAGA 3752
Db 3673 CCTGCC-----TGGGAATAGAGCCAGCCAGGACAGCGTGTGTGACAGATTGGGGCAGG 3724
OY 3753 GAGGAGGGGTGTAGAAATCTCTTACTATCAAGGGAACTGAGTCGTGCACCTGCAAG 3812
Db 3725 ACAGAG-----AGGGAACCTGAGTC-TGGAGCTGGCTAG 3757
OY 3813 TGGATGCTCTCCCTAGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACCTCA 3872
Db 3758 TAGAGCCCTGCTCCAGCTCACAG-----TCCA 3785
OY 3873 GGTCTGAGAGATGAGAGGCTCCCTCAGCCT-GAAGCTATGCAGATAGCCAGGGTTGAA 3931
Db 3786 GGTCTGAGATGTGCGGAGACACTTGCAGCCTGGGCTGTGTGAACAGCAGGCCCTGGTAAA 3845
OY 3932 AGGGGAAGGA-GAGCCTGGATGGGAGCTGTGTGTGGAGGAGGAGGACAGATATTA 3990
Db 3846 TGGGAGAGGGTAGGGGACAGGGAGGAACTGCAGCTGGGGGGGAGGGGACAAAGCATTC 3905
OY 3991 AGCCTGGAAGAGAGGTGACCTTACCAGTTGT--TCAACTCACCTTCAGATTAAAAA 4048
Db 3906 GTCTATATGAAGGTGACCTTACCAGTTGTCTACCCAGTTGTCTCACTCACCTTCAGTTAAAA 3965
OY 4049 TAATGAGGTAAAGGCTCTGGGTAGGGAGGTGGTGT----GAGAGCTCTCTGTCTCTCT 4104
Db 3966 TAACCGAGTAAAGSCCATGTGTGGGGGAGGTGGTGGAGAGGTCTGTCTTCCCA 4025
OY 4105 CTGCATGCTGAGGCCCTTTGGGGAGGAGGATGTGCCAAGGACTTAAAAAAGGCCAT 4164
Db 4026 CTATCTGCCCATCAGCGCTCTGGAGGGGCGGAATGTCTCAAGGACTTAAAAAAGGCCCT 4085
OY 4165 GGAGCCAGAGGGGAGGAGGCAACAGACCTTTTCATGGGCAACCTTTGGGGCCC 4216
Db 4086 GGAGCCGAGGGGCTGGGGGAGCAGACACCTTTTCATGGGCAAAATCTGGGGGGCCC 4137
```

```
RESULT 7
US-09-874-389-7
; Sequence 7, Application US/09874389
; Patent No. US20020152489A1
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; Gossen, Manfred
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible
; Transcription
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/874,389
; FILING DATE: 26-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/161,902
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/275,876
; FILING DATE: 15-JULY-1994
; APPLICATION NUMBER: US 08/270,637
; FILING DATE: 01-JULY-94
; APPLICATION NUMBER: US 08/260,452
; FILING DATE: 14-JUNE-1994
; APPLICATION NUMBER: US 08/076,327
; FILING DATE: 14-JUNE-1993
; APPLICATION NUMBER: US 08/076,726
; FILING DATE: 14-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeCorti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-009CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-874-389-7
```

```
Query Match 5.4%; Score 311.8; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.5e-85;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4277 GGAACTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
Db 63 GGATCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122
OY 4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAG 4396
Db 123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAG 182
OY 4397 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
Db 183 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242
OY 4457 TACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTG 4516
Db 243 TACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTG 302
OY 4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
Db 303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362
OY 4577 CGAGCTCGGTACCAG 4591
Db 363 CGAGCTCGGTACCAG 377
```

```
RESULT 8
US-09-921-650-7
; Sequence 7, Application US/09921650
; Publication No. US2003002315A1
; GENERAL INFORMATION:
; APPLICANT: Bujard, Hermann
; Gossen, Manfred
; TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
```

Activator Fusion Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/921,650
FILING DATE: 03-AUG-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/912,650
FILING DATE: 2001-08-03
APPLICATION NUMBER: US 08/485,978
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-921-650-7
Query Match 5.4%; Score 311.8; DB 10; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.5e-85;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4277 GGAAGTGAAGTTCAGTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
DB 63 GGATCCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCAC 122
QY 4337 TCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAG 4396
DB 123 TCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTT 4456
DB 183 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTT 242
QY 4457 TACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGT 4516
DB 243 TACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGT 302
QY 4517 ATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGT 4576

DB 303 ATAGAGAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGT 362
QY 4577 CGAGTCGGTACCAG 4591
DB 363 CGAGTCGGTACCCG 377
RESULT 9
US-09-241-347-7
; Sequence 7, Application US/09241347
; Publication No. US20040003417A1
; GENERAL INFORMATION:
; APPLICANT: Bujard, Manfred
; Gossen, Manfred
; TITLE OF INVENTION: Animals Transgenic for a Tetracycline-
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/241,347
; FILING DATE: 02-Feb-1999
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,814
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/383,754
; FILING DATE: 03-FEB-1995
; APPLICATION NUMBER: US 08/275,876
; FILING DATE: 15-JULY-1994
; APPLICATION NUMBER: US 08/270,637
; FILING DATE: 01-JULY-1994
; APPLICATION NUMBER: US 08/260,452
; FILING DATE: 14-JUNE-1994
; APPLICATION NUMBER: US 08/076,327
; FILING DATE: 14-JUNE-1993
; APPLICATION NUMBER: US 08/076,726
; FILING DATE: 14-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-009CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-241-347-7
Query Match 5.4%; Score 311.8; DB 11; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.5e-85;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4277 GGAAGTGAAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCAC 4336
DB 63 GGATCCGAGTTTACCACCTCCCTATCAGTATAGAGAAAGTGAAGTCGAGTTTACCAC 122

QY 4337 TCCCTATCAGTGTAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAG 4396
DB 123 TCCCTATCAGTGTAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAG 182
QY 4397 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTT 4456
DB 183 AAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTT 242
QY 4457 TACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGT 4516
DB 243 TACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGT 302
QY 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTT 4576
DB 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTT 362
QY 4577 CGAGTCGGTACCAG 4591
DB 363 CGAGTCGGTACCAG 377

RESULT 10

US-10-134-643-17
; Sequence 17, Application US/10134643
; Publication No. US20030113898A1
; GENERAL INFORMATION:
; APPLICANT: OLSEN, JOHN C.
; APPLICANT: MITROPHANOUS, KYRIACOS ANDREOU
; APPLICANT: ROHL, JONATHAN
; APPLICANT: KINGSMAN, ALAN JOHN
; APPLICANT: ELLARD, FIONA MARGARET
; TITLE OF INVENTION: METHODS FOR PRODUCING HIGH TITRE VECTORS AND
; TITLE OF INVENTION: COMPOSITIONS USED IN SUCH METHODS
; FILE REFERENCE: 078883-0148
; CURRENT APPLICATION NUMBER: US/10/134,643
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/287,048
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleic acid EIAV-TRE hybrid LTR
US-10-134-643-17

Query Match 5.4%; Score 311.4; DB 15; Length 990;
Best Local Similarity 99.7%; Pred. No. 4.7e-85;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4279 AACTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTC 4338
DB 156 AACTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTC 215
QY 4339 CCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAA 4398
DB 216 CCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAA 275
QY 4399 AAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTT 4458
DB 276 AAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTT 335
QY 4459 CCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTAT 4518
DB 336 CCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTAT 395
QY 4519 AGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCG 4578
DB 396 AGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCG 455

QY 4579 AGCTCGGTACCAG 4591
DB 456 AGCTCGGTACCAG 468
RESULT 11
US-10-102-143-12/c
; Sequence 12, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meissner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 12
; LENGTH: 4438
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(1885)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pretO/Sagl-HXGPRT
US-10-102-143-12
Query Match 5.4%; Score 311; DB 16; Length 4438;
Best Local Similarity 96.9%; Pred. No. 1.4e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4263 CTAACGGACAGAGGGAACTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAG 4322
DB 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAG 920
QY 4323 GTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCC 4382
DB 919 GTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCC 860
QY 4383 TATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAA 4442
DB 859 TATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAA 800
QY 4443 GTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACC 4502
DB 799 GTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACC 740
QY 4503 ACTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAG 4562
DB 739 ACTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACCTCCCTATCAGTGTATAG 680
QY 4563 AGAAAAGTGAAGTCGAGTCGGTACC 4589
DB 679 AGAAAAGTGAAGTCGAGTCGGTACC 653
RESULT 12
US-10-102-143-10/c
; Sequence 10, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meissner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10

```
; LENGTH: 4479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(1924)
; NAME/KEY: misc feature
; LOCATION: (1102)
; OTHER INFORMATION: n is disclosed as an asterisk
;
; OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag1-GPP
US-10-102-143-10

Query Match          5.4%; Score 311; DB 16; Length 4479;
Best Local Similarity 96.9%; Pred. No. 1.4e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4263 CTAACGGACAGGAGGAACTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 4322
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 920

Qy 4323 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 4382
Db 919 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 860

Qy 4383 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 4442
Db 859 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 800

Qy 4443 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 4502
Db 799 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 740

Qy 4503 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 4562
Db 739 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 680

Qy 4563 AGAAAAGTGAAAAGTCGAGCTCGGTACC 4589
Db 679 AGAAAAGTGAAAAGTCGAGCTCGGTACC 653

RESULT 13
US-10-102-143-8/c
; Sequence 8, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(3787)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1102)
; OTHER INFORMATION: n is disclosed as an asterisk
;
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pTetO7Sag1-MyoA
US-10-102-143-6

Query Match          5.4%; Score 311; DB 16; Length 6346;
Best Local Similarity 96.9%; Pred. No. 1.7e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4263 CTAACGGACAGGAGGAACTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 4322
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 920

Qy 4323 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 4382
Db 919 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 860

Qy 4383 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 4442
Db 859 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 800

Qy 4443 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 4502
Db 799 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 740

Qy 4503 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 4562
Db 739 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 680

Qy 4563 AGAAAAGTGAAAAGTCGAGCTCGGTACC 4589
Db 679 AGAAAAGTGAAAAGTCGAGCTCGGTACC 653

RESULT 13
US-10-102-143-8/c
; Sequence 8, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1270)..(2001)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag4-GPP
US-10-102-143-8

Query Match          5.4%; Score 311; DB 16; Length 4556;
Best Local Similarity 96.9%; Pred. No. 1.4e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4263 CTAACGGACAGGAGGAACTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 4322
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 920
```

```
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 920
Qy 4323 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 4382
Db 919 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 860
Qy 4383 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 4442
Db 859 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 800
Qy 4443 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 4502
Db 799 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 740
Qy 4503 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 4562
Db 739 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 680
Qy 4563 AGAAAAGTGAAAAGTCGAGCTCGGTACC 4589
Db 679 AGAAAAGTGAAAAGTCGAGCTCGGTACC 653

RESULT 14
US-10-102-143-6/c
; Sequence 6, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(3787)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1102)
; OTHER INFORMATION: n is disclosed as an asterisk
;
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pTetO7Sag1-MyoA
US-10-102-143-6

Query Match          5.4%; Score 311; DB 16; Length 6346;
Best Local Similarity 96.9%; Pred. No. 1.7e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4263 CTAACGGACAGGAGGAACTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 4322
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAA 920

Qy 4323 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 4382
Db 919 GTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCC 860

Qy 4383 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 4442
Db 859 TATCAGTCATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAA 800
Qy 4443 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 4502
Db 799 GTGAAAGTCGAGTTTACCACCTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACC 740
Qy 4503 ACTCCCTATCAGTGATAGAGAAAAAGTGAAAAGTCGAGTTTACCACCTCCCTATCAGTGATAG 4562
```

Db 739 ACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATAG 680
Qy 4563 AGAAAAGTGAAGTCGAGTCGGTACC 4589
Db 679 AGAAAAGTGAAGTCGAGTCGGTACC 653

RESULT 15

US-10-102-143-4/c
; Sequence 4, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meissner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 6423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1270)..(3864)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pTetO7Sag4-MyoA
US-10-102-143-4

Query Match 5.4%; Score 311; DB 16; Length 6423;
Best Local Similarity 96.9%; Pred. No. 1.7e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 4263 CTAACGGACAGAGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAA 4322
Db 979 CTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAA 920
Qy 4323 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCC 4382
Db 919 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCC 860
Qy 4383 TATCAGTCATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 4442
Db 859 TATCAGTCATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 800
Qy 4443 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACC 4502
Db 799 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACC 740
Qy 4503 ACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4562
Db 739 ACTCCCTATCAGTGATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 680
Qy 4563 AGAAAAGTGAAGTCGAGTCGGTACC 4589
Db 679 AGAAAAGTGAAGTCGAGTCGGTACC 653

Search completed: April 25, 2005, 15:35:15
Job time : 2159.75 secs

THIS PAGE BLANK (USPTO)